Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна

ВІСНИК

Харківського національного університету імені В.Н. Каразіна

Серія
«Математичне моделювання.
Інформаційні технології.
Автоматизовані системи управління»

Випуск 69

Серія заснована 2003 р.

BULLETIN

of V.N. Karazin Kharkiv National University

Series
«Mathematical Modeling.
Information Technology.
Automated Control Systems»

Issue 69

First published in 2003

Харків
2026

Засновник журналу Харківський національний університет імені В. Н. Каразіна, Харків, Україна. Рік заснування 2003. Періодичність: 4 випуски на рік. https://periodicals.karazin.ua/mia

Статті містять дослідження у галузі математичного моделювання та обчислювальних методів, інформаційних технологій, захисту інформації. Висвітлюються нові математичні методи дослідження та керування фізичними, технічними та інформаційними процесами, дослідження з програмування та комп'ютерного моделювання в наукоємних технологіях.

Для викладачів, наукових працівників, аспірантів, працюючих у відповідних або суміжних напрямках.

Наказом Міністерства освіти і науки України від 17.03.2020 № 409 наукове фахове періодичне видання Вісник Харківського національного університету імені В.Н. Каразіна серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління» включено до Категорії «Б» Переліку наукових фахових видань України за наступними спеціальностями: 113 – Прикладна математика;122 – Комп'ютерні науки та інформаційні технології; 123 – Комп'ютерна інженерія; 125 – Кібербезпека.

Затверджено до друку рішенням Вченої ради Харківського національного університету імені В. Н. Каразіна (протокол № 5 від 30.03.2026 р.)

Редакційна колегія:

Адреса редакційної колегії: 61022, м. Харків, майдан Свободи, 6, Харківський національний університет імені В. Н. Каразіна, к. 534.
Тел. +380 (57) 705-42-81, Email: journal-mia@karazin.ua.

Мова публікації: українська, англійська.

Статті пройшли внутрішнє та зовнішнє рецензування.

Ідентифікатор медіа у Реєстрі суб'єктів у сфері медіа: R30-04456
(Рішення № 1538 від 09.05.2024 р Національної ради України з питань телебачення і радіомовлення. Протокол № 15)

The founder of the Journal is V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.
Year of foundation 2003. The journal is published four times a year.
https://periodicals.karazin.ua/mia

The articles are present research in the field of mathematical modeling and computing methods, information technologies, information security. New mathematical methods of research and management of physical, technical and information processes, research on programming and computer modeling in science-intensive technologies are covered.

For teachers, researchers, graduate students working in relevant or related fields.

By the order of the Ministry of Education and Science of Ukraine from 17.03.2020 № 409 scientific professional periodical Bulletin of V.N. Karazin Kharkiv National University series "Mathematical modeling. Information Technologies. Automated control systems" is included in Category "B" of the List of scientific professional publications of Ukraine in the following specialties: 113 – Applied Mathematics, 122 – Computer Science and Information Technology; 123 – Computer engineering; 125 – Cybersecurity.

Approved for publication by the decision of the Academic Council of V.N. Karazin Kharkiv National University (Minutes № 5 of 30.03.2026).

Editorial Address: 61022, Kharkiv, Svobodi sq., 6, V.N. Karazin Kharkiv National University, r. 534.
Phone. +380 (57) 705-42-81, Email: journal-mia@karazin.ua.

Language of publication: Ukrainian, English.

The articles pass internal and external review.

Media identifier in the Register of the field of Media Entities: R30-04456
(Decision № 1538 dated May 9, 2024 of the National Council of Television and Radio Broadcasting of Ukraine, Protocol № 15)

ЗМІСТ

CONTENTS

DOI: https://doi.org/10.26565/2304-6201-2026-69-01

УДК (UDC) 004.85:004.056

Bazilevych Kseniia *Associate Professor of Department of Mathematical Modeling and Artificial Intelligence; National Aerospace University "Kharkiv Aviation Institute", Vadym Manko St., 17, Kharkiv, Ukraine 61070*
e-mail: k.bazilevych@khai.edu
https://orcid.org/0000-0001-5332-9545

Parfeniuk Yurii *senior lecturer of Department of Theoretical and Applied Informatics; Karazin Kharkiv National University, Svobody Sq. 4, Kharkiv, Ukraine,61022*
e-mail: parfeniuk@karazin.ua
https://orcid.org/0000-0001-5357-1868

Application of Exploratory Data Analysis for Investigating Factors Influencing Sleep Quality

Relevance. The research of the multifactorial nature of sleep quality requires the analysis of large datasets, which is impossible without the use of exploratory data analysis (EDA) methods to identify hidden patterns. In this regard, the development of approaches for the intelligent analysis of factors influencing sleep is a relevant scientific and technical task. **Goal**. To examine and identify the relationships between physiological, behavioral, and environmental factors and sleep quality using exploratory data analysis methods. **Research methods**. The research was based on exploratory data analysis (EDA) methods, primarily aimed at examining the presence of correlations between sleep quality and variables such as sleep duration, stress level, and physical activity. The subsequent construction of a heatmap was necessary to identify latent relationships and to extract the most relevant features. In addition, a linear regression model, a decision tree model, and a logistic regression model were employed to investigate the factors influencing human sleep quality. **The results**. The results obtained using the developed software application with a graphical user interface for analyzing factors influencing human sleep quality are presented. The software application enables data loading, exploratory data analysis, model construction, and result visualization in a user-friendly format. It supports the application of both classification and regression algorithms, allowing it to be adapted to a wide range of analytical tasks. An analysis of the obtained results was conducted, and models with the highest accuracy, adaptability to complex relationships, and interpretability were identified. **Conclusions**. The obtained results confirm the versatility of decision tree methods for the analysis of sleep-related factors. Their accuracy and algorithmic transparency make this approach optimal for modeling complex interrelationships within the scope of the study. Overall, the analysis of factors influencing sleep using EDA methods enables the transformation of complex data into meaningful analytical models, which represents a relevant task for digital medicine.

Keywords: *Sleep, sleep quality, machine learning, regression, classification, logistic regression, decision tree, eda, python, sleep health dataset*

Як цитувати: Bazilevych K., Parfeniuk Y. Application of Exploratory Data Analysis for Investigating Factors Influencing Sleep Quality. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2025. вип. 69. С.6-19. https://doi.org/10.26565/2304-6201-2026-69-01

How to quote: K. Bazilevych, and Y. Parfeniuk, “Application of Exploratory Data Analysis for Investigating Factors Influencing Sleep Quality”, *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 69, pp. 6-19, 2025. https://doi.org/10.26565/2304-6201-2026-69-01

1. Introduction

Under contemporary conditions of increased psychological and emotional stress, an unstable socio-economic environment, and prolonged exposure to stressors, the issue of sleep quality has become particularly significant. Chronic sleep deprivation and disturbances in sleep architecture adversely affect an individual’s physical, mental, and cognitive functioning, increasing the risk of cardiovascular, endocrine, and mental disorders. Moreover, reduced sleep quality is associated with decreased productivity, impaired attention and memory, as well as elevated levels of anxiety. [1]. Sleep quality is regarded as a complex multifactorial characteristic shaped by the influence of a range of physiological,

psychological, and behavioral factors. These factors include age, body mass index, level of physical activity, stress level, blood pressure, sleep duration, as well as the presence of chronic diseases or sleep disorders such as insomnia or sleep apnea [2]. In countries where societies are exposed to unusual stressful conditions, there is an observed increase in the incidence of insomnia, nocturnal awakenings, difficulties falling asleep, and elevated anxiety levels, which negatively affect the overall health of the population [3]. In Ukraine, according to sociological surveys, nearly half of the population reports a decline in sleep quality since the onset of the full-scale invasion, which is associated with elevated levels of anxiety, forced displacement, and an unstable living environment [4, 5].

Sleep quality is traditionally assessed using questionnaires, clinical instruments, or wearable trackers. However, these methods have certain limitations, ranging from subjectivity to the high cost of equipment, which reduces the accessibility of comprehensive diagnostics for a wide range of users. Consequently, in recent years, there has been growing interest in the application of analytical approaches and machine learning algorithms to study and predict sleep quality based on available physiological and behavioral parameters.

Regression and classification methods allow for the identification of key factors affecting sleep, as well as the construction of predictive models capable of detecting potential disorders or forecasting sleep quality levels [6]. Their application provides flexibility, adaptability to various types of data, and high accuracy, provided that proper approaches to input data processing are employed. Research findings indicate the effectiveness of such models in addressing tasks related both to sleep quality prediction and insomnia diagnosis [7, 8]. The relevance of this research is driven by the need to develop tools that enable the effective analysis of factors influencing sleep using open data and mathematical models. This is particularly valuable in the context of psychoprophylaxis, early risk detection, and the promotion of population mental health.

2. Objective of the study and research tasks

The primary objective of this study is to investigate and identify the relationships between physiological, behavioral, and environmental factors and sleep quality using exploratory data analysis (EDA).

To achieve this objective, the following research tasks are defined:

1. To analyze the characteristics of studying factors influencing human sleep quality.
2. To conduct an analytical review of machine learning methods relevant to achieving the study objectives.
3. To perform exploratory data analysis on the datasets considered in the study.
4. To develop algorithmic models for investigating factors affecting human sleep quality using machine learning techniques.
5. To implement a software application for conducting the study, including visualization of the obtained results.
6. To assess the obtained results.

Object of the Study: The process of investigating factors influencing human sleep quality.

Subject of the Study: Applying Machine Learning Techniques to the Analysis of Factors Affecting Sleep Quality.

The analysis of factors influencing sleep quality requires the completion of the following tasks:

Prediction of the numerical value of a respondent's sleep quality score (ranging from 1 to 10) based on parameters such as age, sleep duration, stress level, and physical activity;

Determination of the presence or absence of a sleep disorder in an individual based on their physiological and behavioral characteristics. Classification is performed by dividing the data into two categories: 'sleep disorder present' and 'sleep disorder absent.

3. Research methods

3.1 Exploratory data analysis

Exploratory data analysis (EDA) is a fundamental stage in the data analytics lifecycle, aimed at thoroughly familiarizing oneself with the available dataset prior to the construction of machine learning models or statistical hypotheses. The concept of EDA was first formulated by the American mathematician and statistician John Tukey in the 1970s. In his seminal work, he emphasized the importance of studying data in their "raw" form to uncover hidden patterns, rather than merely confirming pre-established assumptions [9]. In practice, EDA serves as an intermediate bridge between the data

acquisition stage and data preparation for modeling. Its primary objective is to examine the data structure, identify outliers, missing values, anomalies, potential relationships between variables, and possibly erroneous or incorrect observations. The quality of EDA directly influences both the accuracy of the constructed models and the validity of the analytical decisions made [10].

Within a typical Data Science lifecycle, EDA is usually conducted after the data cleaning stage and prior to the construction of predictive or classification models. Its outcomes can significantly influence the selection of variables, determine the appropriateness of transformations, or reveal new relationships that were not initially apparent. In particular, identifying strong correlations between predictors or visualizing their distributions enables the formulation of valid hypotheses regarding causal relationships [11].

Structurally, EDA encompasses a range of approaches: descriptive statistics (mean, median, variance), graphical visualization (boxplots, histograms, scatter plots), and basic tools for identifying relationships (correlation matrices, cluster analysis). In most modern approaches, EDA results serve not merely as an auxiliary tool but as a full-fledged analytical component that informs the subsequent strategy for model development or decision-making. Thus, EDA is not simply a preparatory stage but a comprehensive analytical practice that enables a researcher to interact with data in an informed manner. Its systematic application helps to avoid critical errors in subsequent stages, improves model quality, and fosters a deeper understanding of the subject domain.

In modern data science, the process of information analysis is implemented as a cycle comprising several sequential stages: data collection, cleaning and preparation, exploratory data analysis, model construction, evaluation of results, implementation, and subsequent monitoring [12]. Within this sequence, the EDA stage functions as a bridge between preliminary data processing and formal modeling, allowing for a deeper understanding of the nature of the data, as well as an assessment of its quality, structure, and statistical patterns.

EDA enables the identification of trends, anomalies, missing values, and multicollinearity, as well as the formulation of hypotheses regarding potential relationships between variables. For this reason, performing EDA is a necessary prerequisite for making informed decisions about the choice of an appropriate machine learning algorithm, the method of data normalization, or the feature engineering strategy [13].

3.2 Methods for studying factors affecting human sleep quality

The investigation of factors influencing human sleep quality is a complex interdisciplinary task that integrates medical, psychological, and analytical aspects. Sleep quality is shaped by a set of interrelated variables: physiological (age, body mass index, heart rate, blood pressure), psychological (stress level, anxiety, depression), and behavioral (physical activity, number of steps, sleep duration and latency, daily routine). In real-world conditions, these variables can exert both independent and combined effects, which significantly complicates the construction of a definitive analytical model.

The complexity of the analysis is further heightened by the high variability of individual characteristics: for example, age and sex may modify the impact of stress on sleep, while physical activity can either improve sleep or worsen it in the case of excessive exertion. Moreover, a significant portion of the variables in sleep quality studies are latent (i.e., not directly observable) and require indirect assessment methods, such as questionnaires, biometric sensors, or psychophysiological testing [2].

Traditionally, scientific practice employs descriptive statistics, correlation analysis, regression, classification, and factor analysis methods to study such complex systems. For example, research on the relationship between BMI and the frequency of nighttime awakenings typically begins with describing mean values within groups and formulating hypotheses regarding their dependence. Subsequently, multivariate analysis methods are applied, allowing the simultaneous consideration of the effects of multiple variables on sleep quality [6].

In contemporary conditions, with the availability of large volumes of data from sleep trackers, questionnaires, and medical devices, machine learning algorithms are increasingly used, allowing for the consideration of nonlinear relationships between parameters. This enables researchers not only to confirm the influence of individual factors but also to develop predictive models of sleep quality for specific population groups. Thus, the study of sleep quality factors requires an integrated approach that combines classical statistical methods with modern data-driven algorithms. This approach allows for the consideration of variable interdependencies, improves diagnostic accuracy, and opens prospects for personalized interventions in the field of sleep medicine.

One of the key directions in the study of factors affecting sleep quality is regression analysis—a classical statistical approach that enables modeling relationships between variables in the form of functional dependencies. In the context of sleep research, this approach makes it possible to predict quantitative characteristics, such as sleep quality scores or sleep duration, based on the influence of predictors including age, stress level, physical activity, body mass index (BMI), and heart rate.

The simplest method is linear regression, which assumes a linear relationship between independent variables and the target variable. In sleep studies, linear regression is often used to assess the extent to which a change in one factor (e.g., the number of daily steps) is associated with an improvement or deterioration in sleep [8]. However, when relationships are more complex or involve interactions among multiple factors, multivariate regression is applied, allowing the simultaneous modeling of the effects of several variables. Furthermore, in the field of behavioral sleep medicine, regression models are used not only for prediction but also to evaluate the importance of individual variables. For example, in the study by Lundgren O., Moneta G. B. (2011) the relationships between depression, anxiety, and physical symptoms (headache, somatic complaints) and subjective sleep quality were analyzed. It was found that depressive symptoms and anxiety were the strongest predictors of poor sleep quality [14]. In the study by Lemma et al. (2012), predictors of poor sleep quality were identified, including stress, anxiety, depression, excessive use of electronic devices, and poor sleep hygiene [15]. The study demonstrates a strong association between psychological state and subjective sleep quality. Such approaches enable a more targeted strategy for addressing factors that contribute to sleep disturbances.

Another popular method is regression trees (decision tree regressors), which offer advantages in interpretability and the ability to account for nonlinear relationships. Unlike linear regression, trees split the data into subgroups based on specific features (e.g., age > 45), allowing for more precise identification of patterns within different respondent groups. The application of regression tree methods in sleep research has practical significance: the resulting models enable the development of automated systems for predicting sleep quality and timely risk detection. When combined with EDA methods, this approach enhances the accuracy and adaptability of solutions in the field of medical technologies.

When the objective of a study is to identify the presence or absence of a sleep disorder, classification methods are employed—a machine learning approach that allows the dataset to be divided into discrete categories. In this case, the target variable is binary or categorical (e.g., 0 - no disorder, 1 - presence of a sleep disorder), while the independent variables consist of physiological, behavioral, or psycho-emotional factors.

One of the fundamental tools of classification is logistic regression—a mathematical model that estimates the probability of belonging to a particular class based on the logistic function. In sleep quality research, logistic regression is used to identify the factors that most strongly influence the likelihood of disorders such as insomnia or sleep apnea [9]. This method also allows for the calculation of the weight of each factor, enabling not only prediction but also the interpretation of the contribution of each feature.

Another popular approach is decision tree classification, which hierarchically splits the dataset based on features, forming rules such as: “if stress level > 6 and sleep duration < 5 hours, then the probability of insomnia is high.” The advantage of this method lies in its interpretability—the researcher can easily trace which factors were decisive in assigning an observation to a particular class. [16].

For tasks with imbalanced classes (for example, when the majority of respondents do not have disorders and only a small portion do), more advanced algorithms such as Random Forest or Support Vector Machines can be applied. These methods improve classification accuracy by simultaneously accounting for multiple factors [17].

Classification models are indispensable in the field of sleep research, as they enable the automated and highly accurate identification of at-risk groups. This is particularly relevant in large population studies or in the development of personalized health monitoring systems.

In contemporary sleep quality research, combining exploratory data analysis (EDA) with the development of predictive or classification models is considered an effective approach. This methodology not only allows for the description of the structure of the available data but also facilitates a deep understanding of the relationships between variables and the identification of hidden patterns, which can subsequently serve as the foundation for machine learning models.

EDA serves as the initial stage—visualizing the distributions of sleep duration, stress level, BMI, heart rate, and other factors allows for the formulation of hypotheses regarding their influence on sleep quality. Using histograms, boxplots, or heatmaps, one can identify, for example, a negative correlation between stress level and sleep quality, or excessive variability in physical activity across different age groups.

Based on such insights, predictive models are developed, including regression models for estimating numerical indicators (e.g., sleep quality scores from 1 to 10) and classification models for determining the likelihood of disorders (insomnia, sleep apnea, etc.). Thus, models are not constructed "blindly" but rely on clearly identified relevant variables revealed through EDA. The integration of EDA and machine learning enhances the interpretability of results and provides a better understanding of causal relationships, which is particularly important in sensitive domains such as sleep research, where an imperfect model may lead to incorrect interpretations of medical risks [18].

Thus, the combination of descriptive, visual, and modeling methods creates an integrated analytical framework that not only enables predictions to be made but also allows them to be interpreted within an applied context.

One of the fundamental tools is linear regression, which models the relationship between predictors (age, sleep duration, stress level) and the sleep quality score. Such a model is easily interpretable, enables the evaluation of each variable's contribution, and helps identify the main factors affecting nighttime rest. At the same time, in cases where nonlinear relationships or a high degree of interdependence among predictors (multicollinearity) are observed, regression trees (Decision Tree Regressors) are preferred. This method constructs the model as a sequence of conditions, allowing for the identification of complex interdependencies between variables without the need for prior transformation.

In the context of sleep research, regression models are often used to estimate the average level of sleep quality based on physiological and behavioral characteristics, such as the frequency of physical activity, the presence of daytime stress, or the average duration of awakenings. Additionally, regression models are employed to assess the effectiveness of corrective interventions—such as changes in sleep hygiene, physical activity, or reduction of stressors. In this way, regression serves not only as an analytical tool but also as a means of monitoring quality of life, particularly within clinical diagnostics or studies of psycho-emotional state.

Another effective method is the decision tree classifier, which constructs a model as a sequence of branching conditions based on predictor values. This structure allows researchers not only to classify subjects but also to trace the logic of decision-making, which is especially valuable in medical or psychological contexts. In more complex cases, where data contain a large number of variables or exhibit high levels of noise, ensemble methods such as Random Forest or Gradient Boosting are applied. These algorithms combine the advantages of multiple decision trees, enhancing result stability and reducing the risk of overfitting. Studies demonstrate that such methods allow for highly accurate classification of sleep disorders even when only a limited number of variables are available [19-22].

4. Research Results

4.1 Initial Data

The study utilized a structured dataset, referred to as the 'Sleep Health and Lifestyle Dataset' [23], which comprises information on individuals' physiological, behavioral, and social characteristics potentially affecting sleep quality. This dataset serves as a well-structured source of input information for the developed software application. The data were stored in a tabular format (.csv), ensuring ease of processing and compatibility with the selected tools in the Python programming language.

The table contains 374 rows, each representing an individual respondent, and twelve variables that are important for modeling purposes. Below is a list of the main columns included in the input data:

1. Person ID - a unique identifier for each study participant. This variable serves an administrative purpose and is not considered in subsequent analysis.
2. Occupation - type of professional employment. A categorical variable reflecting social status and lifestyle.
3. Gender - the respondent's sex (Male or Female). A nominal variable.
4. Age - age of the respondent (quantitative variable).
5. Sleep Duration - number of hours of sleep per day (float).
6. Quality of Sleep - subjective assessment of sleep quality on a scale from 1 to 10. This variable serves as the target in the regression task.
7. Physical Activity Level - level of physical activity, represented as a quantitative value from 0 to 100. Reflects the respondent's overall daily movement, with higher values corresponding to more active individuals. Used as a quantitative predictor in modeling.
8. Stress Level - the respondent's stress level on a scale from 1 to 10 (self-reported).
9. BMI Category - body mass index category (Underweight, Normal, Overweight, Obese).

10. Blood Pressure - arterial blood pressure, presented in the “systolic/diastolic” format, e.g., “120/80.”
11. Heart Rate - heart rate (beats per minute).
12. Daily Steps - average number of steps per day.
13. Sleep Disorder - type of sleep disorder, if present. Possible values: “None,” “Insomnia,” or “Apnea.” This variable serves as the target in the classification task.

For the purpose of further processing, these variables were classified into quantitative, categorical, and target types. Specifically, the variable “Quality of Sleep” serves as the target in the regression task, while “Sleep Disorder” serves as the target in the classification task. The remaining parameters are used as predictors.

During the initial analysis stage, a verification of missing values, variable types, and a visual assessment of proper formatting was conducted. Before proceeding to model construction, each variable was checked for missing or anomalous values, adherence to realistic ranges within the context of the study, and the presence of excessive repetition. Basic visualization methods—histograms, boxplots, and scatter plots—were used to better analyze the distribution of values and to identify potential anomalies. Additionally, the degree of differentiation among variables was assessed to avoid situations in which two features are nearly identical, which could affect model stability. This analysis ensured that the selected factors genuinely influence sleep quality and can be useful for training algorithms.

4.2 Original Dataset

After completing the preliminary data processing, the software automatically generates modeling results, which are presented in an interpretable format—numerical, graphical, or combined. The type of output depends on the chosen task: regression or classification.

When the selected task involves predicting sleep quality, i.e., a numerical indicator representing the respondent’s subjective assessment of sleep, the model produces a point estimate on a scale from 1 to 10. This prediction is based on a combination of input variables, including age, sleep duration, stress level, and physical activity. Such an approach enables analytical forecasting of individual sleep quality prior to direct medical or clinical evaluation. If the user selects a classification task, i.e., determining whether an individual has a sleep disorder, the model analyzes the input features and outputs a categorical result as one of two classes: “sleep disorder present” or “sleep disorder absent.” This approach enables the early identification of potential sleep disorder risks and can be used as an auxiliary tool for preliminary screening.

To evaluate the performance of each model, the software automatically calculates a set of relevant quality metrics. For regression models, MAE (Mean Absolute Error) is used to reflect the average deviation of predicted values from actual values; MSE (Mean Squared Error) assigns greater weight to larger errors; and R^2 (coefficient of determination) indicates the degree of agreement between the model and the actual data. For classification tasks, performance is assessed using metrics such as Accuracy, which measures the overall classification correctness.,

Precision - the accuracy of predicting the positive class; Recall - the completeness of identifying the positive class; F1-score - a combined metric representing the harmonic mean of Precision and Recall.

4.3 Exploratory Data Analysis and Data Preprocessing

The developed application allows EDA to be performed directly through the interface, which implements the construction of four main types of plots. By clicking the “EDA (Data Analysis)” button, the user is presented with graphical visualizations that help form an initial understanding of the data.

The first plot—a boxplot of BMI by sleep disorder type (Figure 1)—illustrates the distribution of body mass index across the groups None, Sleep Apnea, and Insomnia. It can be observed that the mean BMI values differ slightly between groups. For example, respondents with sleep apnea tend to have higher BMI values. The presence of outliers in the plot indicates individual atypical values, which could potentially influence model performance.

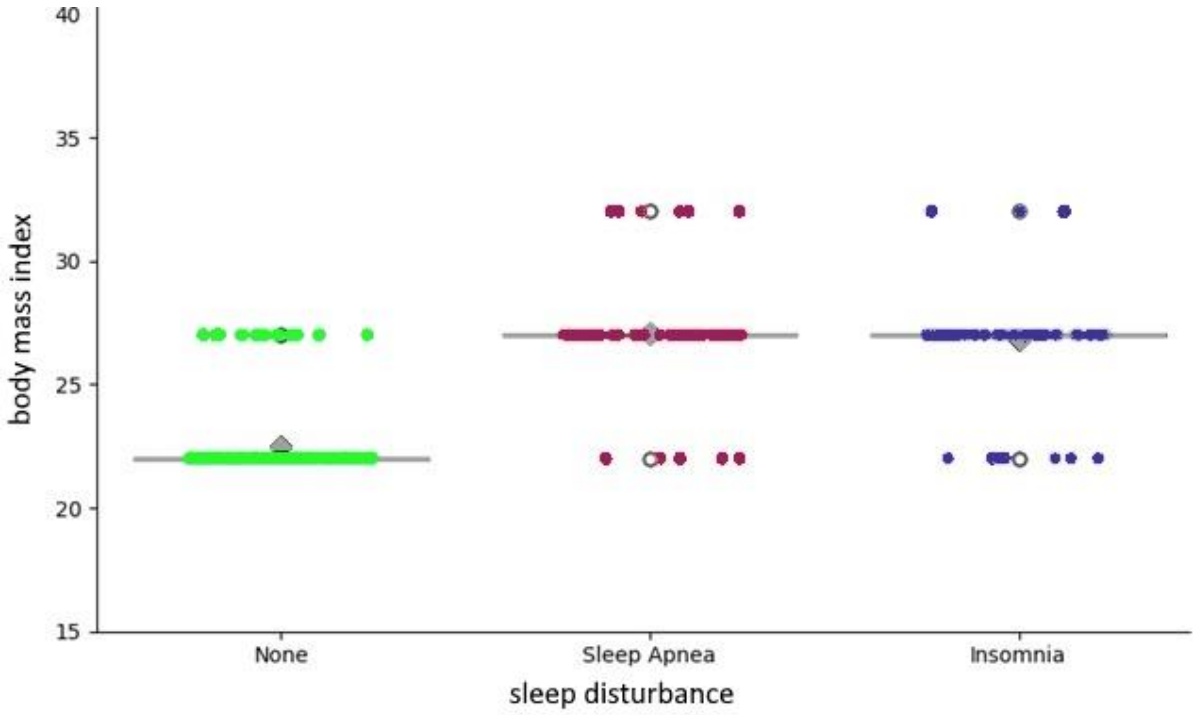


Fig. 1. Boxplot of Body Mass Index (BMI) distribution by sleep disorder type
Рис. 1. Boxplot розподілу індексу маси тіла (BMI) за типами порушень сну

The second plot—a scatterplot showing the relationship between age and sleep duration (Figure 2)—reveals a slight trend: as age increases, the average sleep duration tends to decrease slightly. It is also evident that respondents with different sleep disorders are concentrated in specific ranges. For example, most individuals with insomnia fall within the 30-45-year age range and have sleep durations of less than 7 hours.

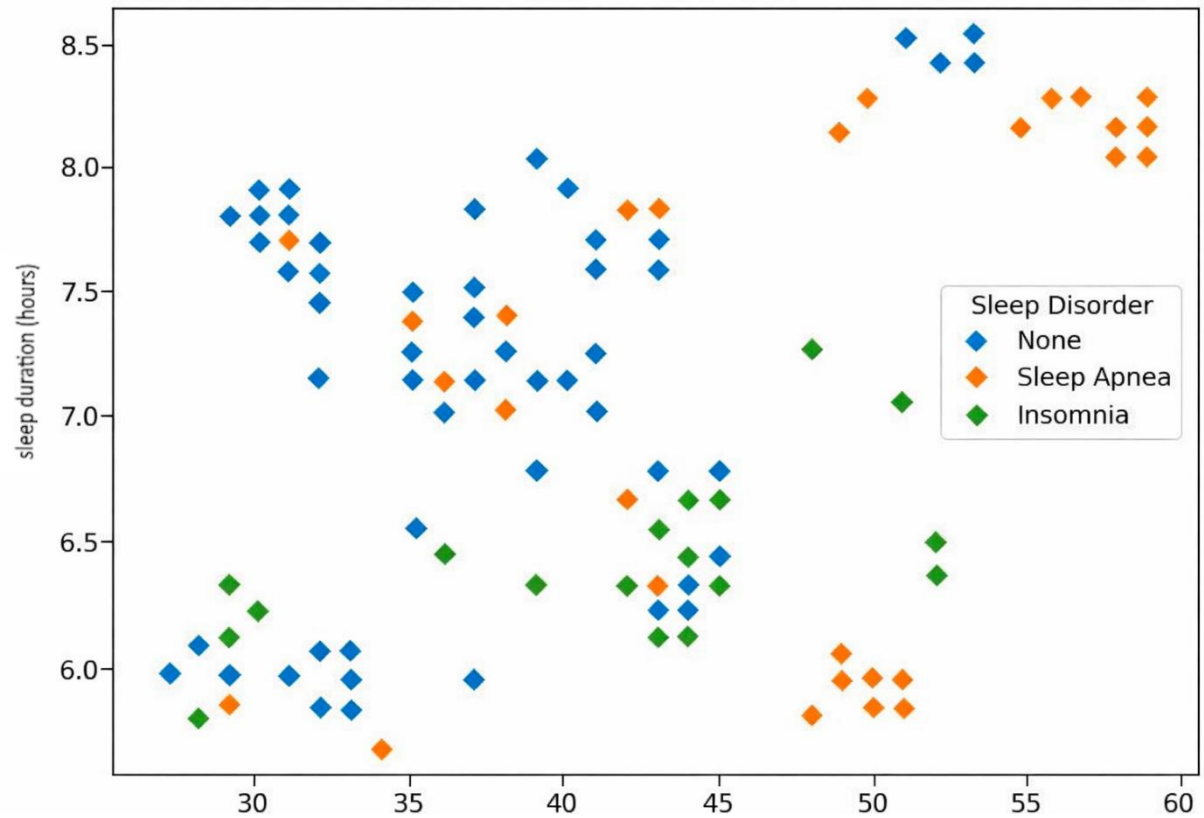


Fig. 2. Scatter plot: relationship between age and sleep duration
Рис. 2. Діаграма розсіювання: взаємозв'язок між віком та тривалістю сну

The third plot—the distribution of sleep disorder types (Figure 3)—illustrates class imbalance. The largest number of respondents belongs to the “None” group, indicating no sleep disorders, whereas the “Sleep Apnea” and “Insomnia” groups are represented by significantly fewer individuals. This distribution should be taken into account when constructing classification models, as the imbalance may lead to bias toward the majority class.

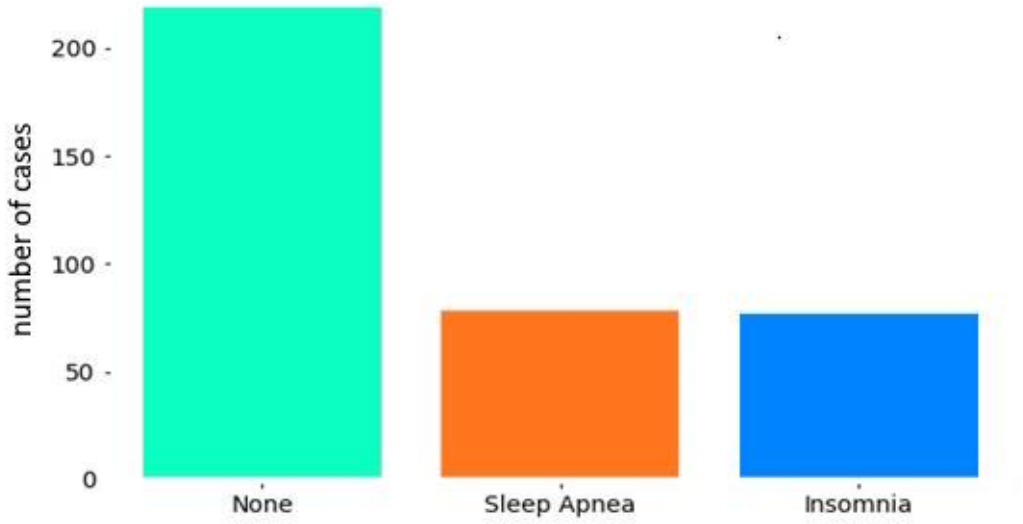


Fig. 3. Histogram of respondents by sleep disorder type
Рис. 3. Гістограма розподілу респондентів за типами порушень сну

The fourth plot - a correlation heatmap (Figure 4) - visually illustrates the relationships between numerical features. The strongest positive correlation is observed between sleep duration and sleep quality (coefficient ≈ 0.88), whereas stress level shows a negative correlation with sleep quality (≈ -0.90). This supports the hypothesis that chronic stress significantly reduces sleep quality. A strong correlation is also observed between daily steps and physical activity level (≈ 0.77), which is consistent with the nature of these indicators. The provided heatmap illustrates how various factors (such as age, sleep, stress, and activity) are interrelated. Red colors indicate that as one variable increases, the other tends to increase as well (positive correlation), while blue colors indicate that as one variable increases, the other tends to decrease (negative correlation). The intensity of the color reflects the strength of the correlation.

Considering the variable “target” (the dependent variable), it shows a moderate positive correlation with Age (coefficient 0.43) and Person ID (0.45). This suggests that, in general, the target value slightly increases with age. The Person ID is almost perfectly correlated with Age (0.99), indicating that older individuals have higher IDs in this dataset. In contrast, Sleep Duration (-0.34) and Sleep Quality (-0.31) exhibit a weak negative correlation with the target, meaning that longer and higher-quality sleep is somewhat associated with lower target values. Heart Rate shows a weak positive correlation with the target (0.33). Meanwhile, Physical Activity Level, Stress Level, and Daily Steps display very weak or nearly nonexistent linear relationships with the target, with coefficients close to zero.

Among other factors, several very strong relationships stand out. The most notable is the extremely strong negative correlation between Sleep Quality and Stress Level (-0.90), indicating that higher stress levels are associated with a marked decrease in sleep quality. Sleep Duration also decreases significantly with increasing stress levels (-0.81). It is logical that Sleep Duration and Sleep Quality are strongly positively correlated (0.88), meaning that longer sleep is generally of higher quality. Additionally, higher Stress Levels are noticeably correlated with higher Heart Rate (0.67), whereas better Sleep Quality and longer Sleep Duration are associated with lower Heart Rate (-0.66 and -0.52, respectively). Physical Activity Level is also strongly correlated with Daily Steps (0.77), which is expected.

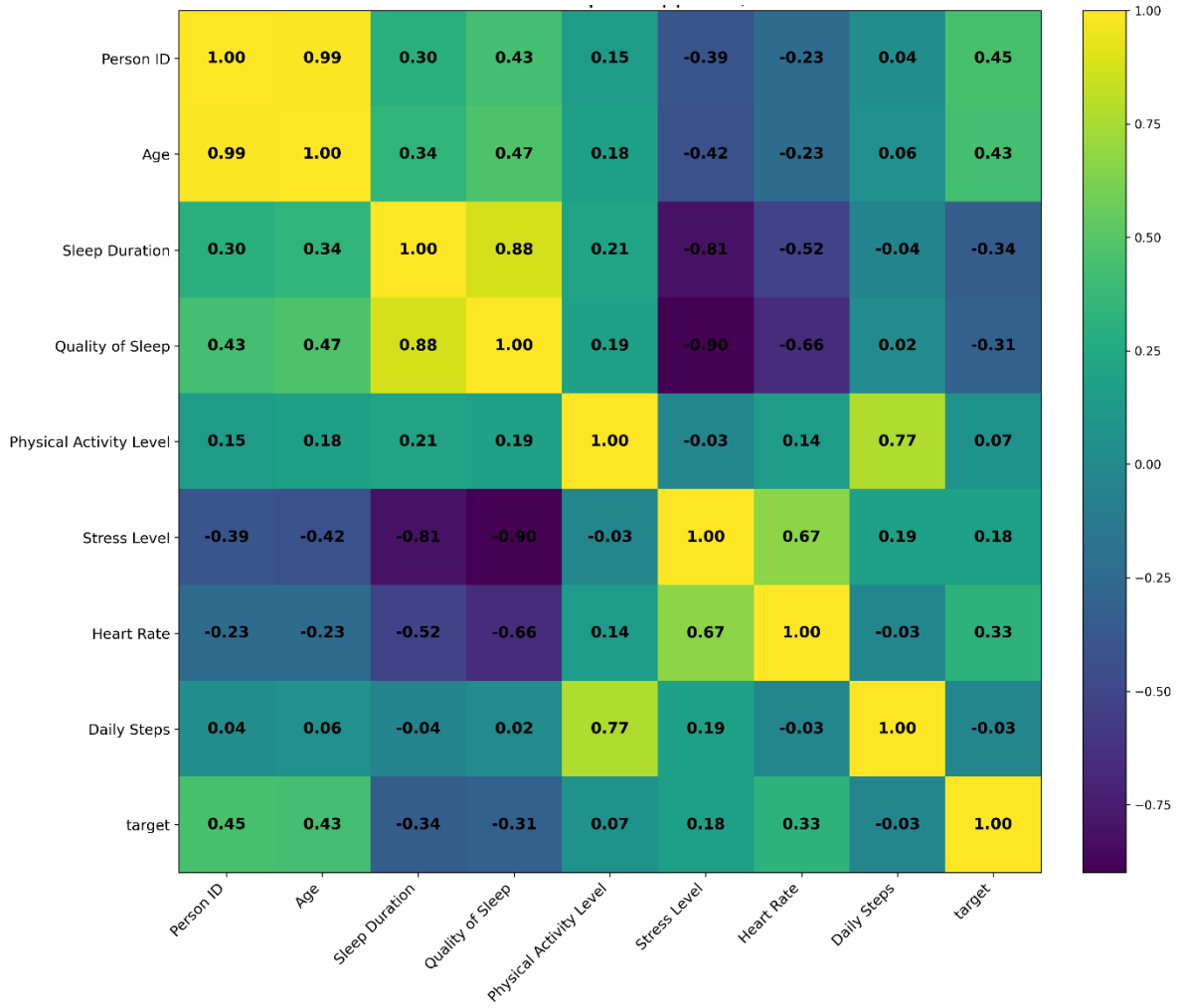


Fig. 4. Теплова карта кореляцій між числовими ознаками
Рис. 4. Теплова карта кореляцій між числовими ознаками

In summary, performing EDA allowed for verification of the dataset quality, identification of key relationships, and formulation of hypotheses regarding significant factors. These findings subsequently serve as the foundation for constructing regression and classification models.

4.4 Analysis of the results obtained

As a result of running the linear regression model, predictions of sleep quality were made based on the variables age, sleep duration, stress level, and physical activity level. The obtained numerical indicators indicate high model quality. The Mean Absolute Error (MAE) is 0.32, reflecting a small average deviation of predicted values from the actual values. The Mean Squared Error (MSE) is 0.18, demonstrating model stability with a low incidence of large errors. The coefficient of determination (R^2) is 0.88, indicating that 88% of the variance in the target variable is explained by the selected predictors.

For a visual representation of the model's accuracy, a scatter plot was constructed (Figure 5), with the true sleep quality values on the X-axis and the predicted values on the Y-axis.

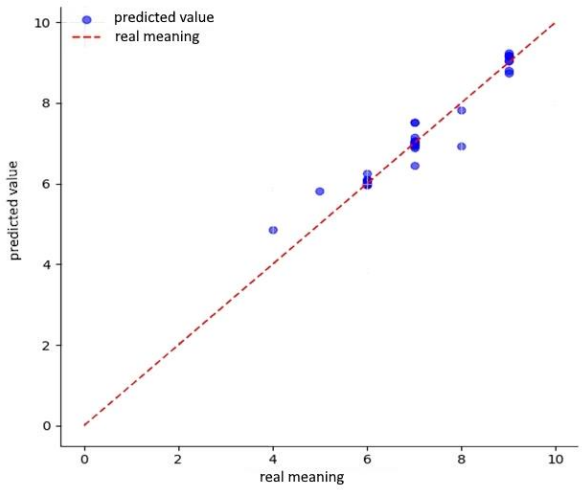


Fig. 5. Scatter plot for the Linear Regression model
Рис. 5. Діаграма розсіювання для моделі Linear Regression

The ideal line is represented by a dashed line and corresponds to a perfect match between predicted and actual values. Analysis of the scatter plot shows that most points are clustered along this line, confirming the adequacy of the model. No visible systematic deviations or outliers are observed.

Thus, the linear regression model demonstrated a high capacity to predict the target variable based on the available parameters, making it a suitable baseline tool for predictive analysis in tasks related to sleep quality assessment.

The next stage involved using a Decision Tree algorithm for the regression task, which allows the construction of an interpretable model based on a hierarchical splitting of the dataset according to feature values. After training the model, the resulting metrics indicate very high prediction accuracy. The MAE is 0.03, reflecting an almost negligible average error. The MSE is also 0.03, indicating minimal deviations in the predictions. The highest performance is demonstrated by the coefficient of determination, R^2 = 0.98, meaning that 98% of the variance in the target variable is explained by the selected features.

Such an R^2 value indicates an almost complete correspondence between predicted and actual values. At the same time, this very high accuracy may suggest a risk of overfitting, especially given the limited size of the training dataset, which necessitates additional validation on an external dataset. It should be noted that the structure of the decision tree allows interpretation of the decision-making process, which can be useful when applying the model in medical or social research.

In the scatter plot (Figure 6), almost all points lie on or very close to the ideal prediction line. Deviations are minimal, so the graph practically demonstrates a match between predicted and actual values. This confirms the high accuracy of the model as well as its potential for further use in sleep quality prediction tasks.

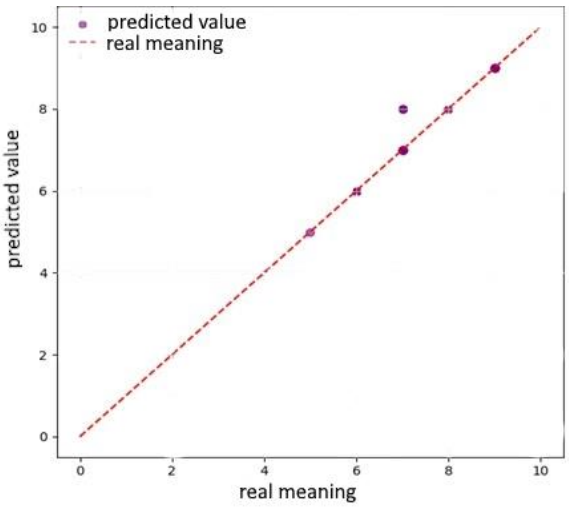


Fig. 6. Scatter plot for the Decision Tree Regressor model
Рис. 6. Діаграма розсіювання для моделі Decision Tree Regressor

Thus, the decision tree model demonstrates high predictive accuracy, making it suitable for tasks where minimizing error, ensuring interpretable decision logic, and adapting to new input data are important. For the classification task, a logistic regression model was used to predict the probability of each respondent belonging to one of two classes: absence of sleep disorders or presence of a sleep disorder. Features included age, stress level, and physical activity. The target variable was formulated as a binary indicator: class “0” - normal sleep, class “1” - presence of a disorder (insomnia or apnea).

The obtained results indicate moderate performance of the model. Accuracy is 0.72, meaning that 72% of predictions were correct. Precision = 0.68 indicates that nearly 7 out of 10 predictions regarding the presence of a sleep disorder were correct. Recall = 0.61 shows that the model identified 61% of all actual sleep disorder cases. F1-score = 0.64 reflects the balance between precision and recall and characterizes the overall classification quality.

Additionally, a histogram of predicted classes was constructed (Figure 7), illustrating the distribution of respondents with and without sleep disorders.

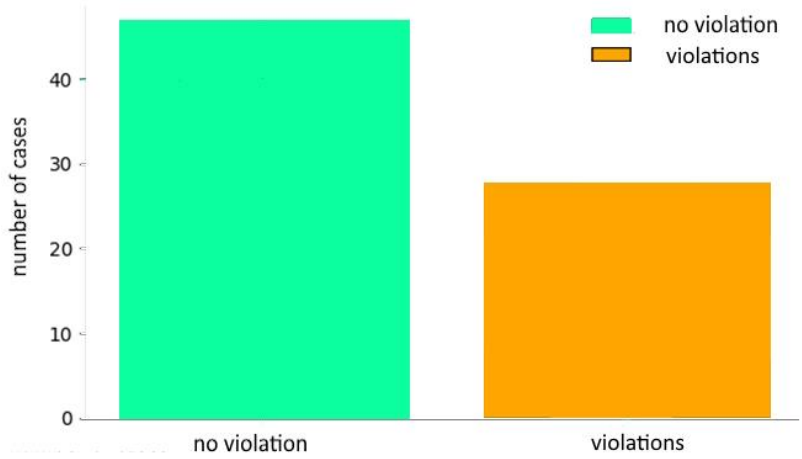


Fig. 7. Predicted class distribution for the Logistic Regression model
Рис. 7. Розподіл передбачених класів для моделі Logistic Regression

The plot shows that the model slightly favors predicting the “no disorder” class, which may be related to the class imbalance in the dataset. However, the overall shape of the distribution corresponds to the actual data structure, indicating that the algorithm behaves adequately.

Thus, the logistic regression model demonstrated satisfactory accuracy and balance in classifying the binary variable. Considering its simplicity, interpretability, and stability, the model can be applied for preliminary screening of sleep disorder risks. Within the binary classification task, a decision tree model was also employed, allowing the creation of a branched classification logic based on the values of the input variables. The input features included age, stress level, and physical activity, while the target variable indicated the presence or absence of a sleep disorder. Unlike logistic regression, the decision tree allows modeling nonlinear relationships and can adapt to more complex data structures.

The obtained numerical metrics indicate a high quality of classification. An Accuracy of 0.89 shows that the model correctly classified 89% of the examples. Precision = 0.87 means that predictions of the “sleep disorder” class were correct in 87% of cases, while Recall = 0.87 indicates that 87% of all true cases of sleep disorders were detected. The F1-score = 0.87 reflects a high level of model balance and its ability to handle moderately imbalanced datasets (Figure 4.16).

The histogram (Figure 7) illustrates the distribution of predicted classes after applying the decision tree model for the classification task. The target variable was “Sleep Disorder”, which was previously encoded in binary format. The task was to determine whether a respondent had a sleep disorder based on features such as age, stress level, and physical activity. The plot allows for assessing the balance of predictions between class 0 (no disorder) and class 1 (disorder present) and confirms that the model does not favor one class over the other. The histogram shows that the classification is relatively uniform, without significant bias, which is particularly important when the positive class is underrepresented.

In conclusion, the decision tree model demonstrated the best performance among all classification approaches, making it suitable for practical applications in detecting sleep disorder risks, considering its accuracy, stability, and interpretability. A comparative summary of all models is presented in Table 1.

Table 1. Comparative Performance of Developed Models
Таблиця 1. Порівняльна характеристика результатів побудованих моделей

Модель	Тип задачі	MAE	MSE	R^2	Accuracy	Precision	Recall	F1-score
Linear Regression	Prediction of Subjective Sleep Quality	0.32	0.18	0.88	-	-	-	-
Decision Tree Regressor	Prediction of Subjective Sleep Quality	0.03	0.03	0.98	-	-	-	-
Logistic Regression	Detection of the Presence or Absence of Sleep Disorders	-	-	-	0.72	0.68	0.61	0.64
Decision Tree Classifier	Detection of the Presence or Absence of Sleep Disorders	-	-	-	0.89	0.87	0.87	0.87

5. Conclusions and prospects for further research

As a result of the practical application of the developed software tool, a complete data analysis cycle was carried out - from preliminary examination of the variables to the construction of machine learning models and evaluation of their performance. Each implemented model demonstrated a different level of accuracy, allowing for a comparative assessment of their advantages in the context of the posed tasks. The results of the exploratory data analysis (EDA) confirmed the presence of strong correlations between sleep quality and variables such as sleep duration, stress level, and physical activity. The construction of a heatmap enabled the identification of hidden dependencies and the determination of the most relevant features for further modeling. Among the regression models, the decision tree yielded the best performance, achieving the lowest error values (MAE = 0.03; MSE = 0.03) and the highest coefficient of determination ($R^2 = 0.98$), indicating the model's strong ability to capture patterns in the data. The linear regression model also produced satisfactory results, though it was outperformed by the decision tree in terms of predictive accuracy.

Among the classification models, the decision tree proved to be the most effective, achieving high values across all key metrics (Accuracy, Precision, Recall, F1-score = 0.87). Logistic regression showed moderate performance, which may be attributed to the linear nature of the algorithm and the potential complexity of the data. In summary, the results indicate that the decision tree algorithm is the most suitable for both regression and classification tasks within this subject area. Its high accuracy, adaptability to complex relationships, and interpretability make it an optimal choice for further applications in analyzing factors affecting sleep quality.

Prospects for further research lie in scaling the developed methodology to significantly larger and more heterogeneous datasets, including those collected in real time via Internet of Things (IoT) sensors and wearable devices. This will enable a transition from identifying general patterns to building high-precision models for personalized monitoring, capable of adapting to individual user characteristics. Thanks to the high interpretability of the decision tree algorithm, the obtained results can serve as a foundation for developing intelligent clinical decision support systems in digital healthcare. In the future, this approach will not only allow for the prediction of sleep disorder risks but also facilitate the automatic generation of scientifically grounded lifestyle recommendations to improve the overall psychophysiological well-being of the population. Further application of this methodology to large-scale datasets will contribute to additional model validation and assessment of its robustness against variability in clinical research samples.

REFERENCES

1. Hobson J. A. Sleep is of the brain, by the brain and for the brain. *Nature*. 2005;437(7063):1254–1256. https://doi.org/10.1038/nature04283.
2. Irish L. A., Kline C. E., Gunn H. E., Hall M. H., Buysse D. J. The role of sleep hygiene in promoting public health: a review of empirical evidence. *Sleep Medicine Reviews*. 2015;22:23–36. https://doi.org/10.1016/j.smrv.2014.10.001.
3. Deng Z., Xie L., Wang Y., Li Y., Huang X., Sun L. Application of logistic regression in diagnosis of OSA severity. *Sleep and Breathing*. 2020;24(4):1379–1387. https://doi.org/10.1007/s11325-020-02029-x.
4. Korost Ya. V., Shkvarok A. K. Assessment of sleep quality of the population of Ukraine during martial law and the risk of cardiovascular complaints associated with clinically expressed insomnia. *Clinical and Preventive Medicine*. 2023;(7):68–73. https://doi.org/10.31612/2616-4868.7.2023.09
5. Ukrinform. Half of Ukrainians have sleep problems. URL: https://www.ukrinform.ua/rubric-society/3958411-polovina-ukrainciv-maut-problemi-zi-snom.html (accessed 12.12.2025).).
6. Carskadon M. A., Dement W. C. Normal human sleep: an overview. In: Kryger M. H., Roth T., Dement W. C., editors. *Principles and Practice of Sleep Medicine*. 6th ed. Philadelphia: Elsevier; 2017. p. 15–24.
7. Freeman D., Sheaves B., Waite F., Harvey A. G. Sleep disturbance and psychiatric disorders: a review of meta-analyses. *The Lancet Psychiatry*. 2017;4(8):684–698. https://doi.org/10.1016/S2215-0366(17)30150-9.
8. Khalid M., Klerman E. B., McHill A. W., Phillips A. J. K., Sano A. SleepNet: attention-enhanced robust sleep prediction using dynamic social networks. *arXiv preprint*. 2024;arXiv:2401.11113. Available from: https://arxiv.org/abs/2401.11113 (accessed 12 Dec 2025).
9. Tukey J. W. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley; 1977. 688 p.
10. Kelleher J. D., Tierney B. *Data Science: An Introduction*. Cambridge, MA: MIT Press; 2018. 280 p.
11. Behrens J. T., Yu C. H. Exploratory data analysis. In: Schinka J. A., Velicer W. F., editors. *Handbook of Psychology*. Vol. 2. Thousand Oaks, CA: SAGE; 2003. p. 33–48.
12. Provost F., Fawcett T. *Data Science for Business*. Sebastopol, CA: O'Reilly Media; 2013. 414 p.
13. Dasgupta A. *Practical Data Analysis*. Birmingham: Packt Publishing; 2014. 342 p.
14. Lundgren O., Moneta G. B. Associations of subjective sleep quality with depression, anxiety, and physical symptoms. *Scandinavian Journal of Psychology*. 2011;52(6):544–550. https://doi.org/10.1111/j.1467-9450.2011.00910.x
15. Lemma S., Gelaye B., Berhane Y., Worku A., Williams M. A. Sleep quality and its psychological correlates among university students in Ethiopia: a cross-sectional study. *BMC Psychiatry*. 2012;12:237. https://doi.org/10.1186/1471-244X-12-237
16. Trujillano J., Gil-Sánchez D., Párraga-Martínez I., Flores-Mateo G. Methodological review of classification trees for risk stratification in health research // *Nutrients*. – 2025. – Vol. 17, № 11. – Article 1903. – doi: 10.3390/nu17111903.
17. Rezvani S., Pourpanah F., Lim C. P., Wu Q. M. J. Methods for class-imbalanced learning with support vector machines: a review and an empirical evaluation. *Soft Computing*. 2024;28:11873–11894. https://doi.org/10.1007/s00500-024-09931-5
18. Fu W. Exploratory data analysis and machine learning models for stroke prediction. In: *Proceedings of the 1st International Conference on Data Analysis and Machine Learning (DAML 2023)*; 2024. p. 211–217. https://doi.org/10.5220/0012783300003885.
19. Permana K. E., Iramina K. Enhancing sleep stage classification with single-channel EEG: feature extraction and Random Forest–XGBoost model. *IEEE Access*. 2025;13:149554–149566. https://doi.org/10.1109/ACCESS.2025.3599828.
20. Gao Q., Wu K. Automatic sleep staging based on power spectral density and random forest. *Journal of Biomedical Engineering*. 2023;40(2):280–285, 294. https://doi.org/10.7507/1001-5515.202207047.

21. Wang Y., Ye S., Xu Z., Chu Y., Zhang J., Yu W. Research on sleep staging based on support vector machine and extreme gradient boosting algorithm. *Nature and Science of Sleep*. 2024;16:1827–1847. https://doi.org/10.2147/NSS.S467111.
22. Xu X., Zhang B., Xu T., Tang J. An effective and interpretable sleep stage classification approach using multi-domain EEG and EOG features. *Bioengineering*. 2025;12(3):286. https://doi.org/10.3390/bioengineering12030286.
23. Sleep Health and Lifestyle Dataset. Kaggle. Available from: https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset (accessed 12 Dec 2025).

Базілевич Ксенія Олексіївна *к.т.н., доцент, доцент кафедри математичного моделювання та штучного інтелекту Національного аерокосмічного університету "Харківський авіаційний інститут", вул. Вадима Манька, 17, 61070, Харків*
e-mail: k.bazilevych@khai.edu
https://orcid.org/0000-0001-5332-9545

Парфенюк Юрій Леонідович *PhD, старший викладач, кафедри теоретичної та прикладної інформатики*
Харківський національний університет імені В. Н. Каразіна, майдан Свободи 4, 61022, Харків
e-mail: parfeniuk@karazin.ua
https://orcid.org/0000-0001-5357-1868

Застосування розвідувального аналізу даних для дослідження факторів, що впливають на якість сну

Актуальність. Дослідження багатофакторної природи якості сну потребує аналізу великих масивів даних, що неможливо без застосування методів розвідувального аналізу (EDA) для виявлення прихованих закономірностей. У зв'язку з цим, розробка підходів до інтелектуального дослідження чинників впливу на сон є актуальною науково-технічною задачею **Мета**. Дослідити та виявити взаємозв'язки між наборами фізіологічних, поведінкових та зовнішніх факторів та якістю сну за допомогою exploratory data analysis. **Методи дослідження**. Дослідження базувалось на методах розвідувального аналізу даних (EDA) для того, щоб в першу чергу дослідити наявність кореляцій між якістю сну та такими змінними, як тривалість сну, рівень стресу й фізична активність. Подальша побудова теплової карти необхідна була для виявлення прихованих залежності та отримання найбільш релевантних ознак. Також було використано модель лінійної регресії, модель дерева рішень, модель логістичної регресії для дослідження факторів, що впливають на якість сну людини. **Результати**. Представлено результати, які отримані за допомогою розробленого програмного додатку з графічним інтерфейсом для дослідження факторів, що впливають на якість сну людини. Програмний додаток дозволяє виконувати завантаження даних, проводити розвідувальний аналіз, будувати моделі та виводити результати у зручному форматі, підтримує застосування як класифікаційних, так і регресійних алгоритмів, дозволяючи адаптувати її до різних агналітичних завдань. Було проведено аналіз отриманих результатів та виявлено моделі з найбільшою точністю, адаптивністю до складних зв'язків і пояснюваністю. **Висновки**. Отримані результати підтверджують універсальність методу дерев рішень для аналізу факторів сну. Його точність і прозорість алгоритмів роблять цей підхід оптимальним для моделювання складних взаємозв'язків у межах дослідження. В цілому, аналіз чинників впливу на сон за допомогою методів EDA дозволяє трансформувати складні дані у змістовні аналітичні моделі, що є актуальним завданням для цифрової медицини.

***Ключові слова**: сон, якість сну, машинне навчання, регресія, класифікація, логістична регресія, дерево рішень, eda, python, sleep health dataset.*

Надійшла у першій редакції 11.02.2026; рекомендована до друку 19.03.2026; опублікована 03.04.2026.
The first version has been received on 11.02.2026; was recommended for printing on 19.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-02

УДК (UDC) 004.056.5:004.4

Братко Дмитро Вячеславович *магістрант, курсант*
Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського", вул. Верхньоключова 4, м. Київ, Україна, 03056
e-mail: loading.2285@gmail.com
https://orcid.org/0009-0001-0863-9285

Кубрак Володимир Олександрович *PhD, Старший викладач Спеціальної кафедри № 1*
Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського", вул. Верхньоключова 4, м. Київ, Україна, 03056
e-mail: v.kubrak@kpi.ua;
https://orcid.org/0000-0001-8877-5289

Матійко Александра Андріївна *PhD, Доцент Спеціальної кафедри №1*
Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського", вул. Верхньоключова 4, м. Київ, Україна, 03056
e-mail: alexm1710@ukr.net;
https://orcid.org/0000-0002-6947-5958

Автономна оркестрована система реагування на інциденти на основі SIEM

Актуальність. Сучасні інформаційні системи генерують події безпеки з різних джерел, таких як журналів операційних систем і сервісів, мережевих сенсорів, сканерів вразливостей та інших засобів моніторингу. У таких умовах SIEM дає змогу централізовано збирати, індексувати та корелювати телеметрію, однак сам перехід від результату аналітики до практичного реагування часто залишається недостатньо формалізованим. Це призводить до затримок, залежності від ручних дій, складнощів із повторюваністю процедур і відсутності єдиного механізму підтвердження виконаних реакцій. Додатковою проблемою є безпечний автономний доступ до кінцевих вузлів під час інциденту, коли неприпустимими є як ручні підтвердження SSH-з'єднання, так і небезпечна довіра до першого ключа. У зв'язку з цим актуальною є побудова архітектурного моста між SIEM-аналітикою та системою оркестрації, здатного забезпечити кероване, відтворюване й аудитоване реагування на інциденти незалежно від первинного джерела подій.
Метою роботи є обґрунтування та експериментальна перевірка архітектурного підходу до автономного оркестрованого реагування на інциденти, у межах якого результати аналітики SIEM перетворюються на структурований інцидентний запис і далі використовуються для запуску процедур реагування в системі оркестрації. Для досягнення цієї мети передбачено опис детектора у декларативному вигляді, стандартизацію інцидентного запису, керування повторними спрацюваннями через унікальний ключ інциденту та інтервал блокування повторного запуску, узгодження цільових активів із даними inventory, журналювання результатів виконання, а також реалізацію безпечного доступу до кінцевих вузлів на основі SSH Host CA. Демонстраційним прикладом обрано сценарій виявлення та реагування на SSH brute-force.
Результати. У результаті дослідження сформовано й експериментально перевірено архітектурний підхід, який поєднує SIEM-аналітику з автоматизованим виконанням реакційних дій на цільових активах. Показано, що результат аналітичного запиту в SIEM може бути послідовно перетворений на інцидентний запис, використаний для побудови унікального ключа інциденту, перевірки політик повторного запуску, узгодження активу з inventory та передачі параметрів до playbook. Реалізований прототип підтвердив технічну можливість побудови повного циклу від виявлення події в SIEM до виконання реакційної процедури на кінцевому вузлі та фіксації результату у структурованому журналі. Окремо підтверджено, що використання SSH Host CA дає змогу забезпечити безпечний автономний доступ до кінцевих вузлів без ручного підтвердження під час інциденту. Отримані результати також показали, що запропонована архітектура може бути масштабована на інші сценарії реагування за умови зміни правил виявлення та процедур виконання.
Висновки. Отримані результати підтверджують, що поєднання SIEM-аналітики з системою оркестрації дає змогу реалізувати керований контур автономного реагування на інциденти. Результат аналітики в SIEM перетворюється на інцидентний запис, який використовується для контролю повторних запусків, узгодження цільового активу з inventory та запуску сценарію реагування. Практична перевірка на прикладі SSH brute-force підтвердила технічну здійсненність

такого підходу: реалізовано повний цикл від виявлення події до виконання реакції та фіксації її результату в журналі. Запропонована архітектура придатна для реагування на інциденти, зафіксовані з використанням різних джерел подій, якщо їх результати агрегуються та корелюються в SIEM. Використання SSH Host CA забезпечує безпечний автономний доступ до кінцевих вузлів без ручного підтвердження під час інциденту. Подальший розвиток роботи доцільно пов'язати з програмною реалізацією модуля-міста, розширенням бібліотеки сценаріїв реагування та передаванням журналу реагування до SIEM для подальшого аналізу.
Ключові слова: *Splunk, Ansible, SIEM, оркестроване реагування, автономне реагування, SSH brute-force, SSH Host CA.*

Як цитувати: Братко Д. В., Кубрак В. О., Матійко А. А. "Автономна оркестрована система реагування на інциденти на основі SIEM". *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.20-32. https://doi.org/10.26565/2304-6201-2026-69-02

How to quote: D. Bratko, V. Kubrak, and A. Matiiko, "Autonomous Orchestrated Incident Response System Based on SIEM", Bulletin of V. N. Karazin Kharkiv National University, series Mathematical Modelling. Information Technology. Automated Control Systems, vol. 69, pp. 20–32, 2026. https://doi.org/10.26565/2304-6201-2026-69-02 [in Ukrainian]

Вступ

Стрімке ускладнення сучасних інформаційних інфраструктур і зростання кількості джерел телеметрії суттєво змінюють підходи до виявлення та обробки інцидентів інформаційної безпеки. У практичних середовищах події, що можуть свідчити про інцидент, надходять із журналів операційних систем і сервісів, мережевих сенсорів, сканерів вразливостей, засобів контролю доступу та інших систем моніторингу. За таких умов саме SIEM-платформи дедалі частіше виступають центральним рівнем збору, індексації та кореляції подій, оскільки дозволяють звести різнорідну телеметрію до єдиного аналітичного простору та виявляти підозрілі закономірності на основі узагальнених правил і запитів [1–4]. Водночас саме виявлення інциденту ще не гарантує своєчасного реагування, оскільки на практиці між аналітичним результатом SIEM і фактичним виконанням дій на цільовому активі часто існує розрив.

Ця проблема особливо помітна в середовищах, де реагування повинно бути швидким, повторюваним і незалежним від первинного джерела події. Більшість існуючих контурів автоматизації або орієнтовані на окремий тип телеметрії, або вимагають побудови окремих інтеграцій для кожного нового джерела даних. У результаті організація отримує не єдиний механізм реагування, а набір розрізнених сценаріїв, які складно супроводжувати, масштабувати та перевіряти з позиції аудиту. Крім того, навіть за наявності коректно налаштованих детекторів, автономне реагування ускладнюється необхідністю узгодження ідентифікаторів активів між SIEM та системою оркестрації, контролем повторних запусків однієї й тієї самої процедури, а також потребою фіксації результатів виконаних дій у вигляді доказового журналу.

Окрему проблему становить безпечний доступ до кінцевих вузлів у момент реагування. Якщо система оркестрації під час інциденту залежить від ручного підтвердження SSH-з'єднання, автоматизація втрачає оперативність. Якщо ж довіра до вузла формується через автоматичне прийняття першого ключа, виникає ризик підміни хоста та компрометації всього контуру реагування. Саме тому задача побудови автономного оркестрованого реагування повинна розглядатися не лише як проблема запуску playbook або скриптів, а як комплексна архітектурна задача, що охоплює уніфікацію інцидентного запису, керування повторюваністю дій, узгодження цільових активів і формування безпечного довіреного середовища для виконання процедур реагування.

У цьому контексті доцільним є використання проміжного архітектурного модуля-міста між SIEM та системою оркестрації, який дозволяє перетворити результат аналітики на структурований інцидентний запис, застосувати до нього визначені політики запуску та передати підготовлений набір параметрів до системи реагування. Такий підхід дає змогу побудувати універсальний контур автономного реагування для інцидентів, зафіксованих із використанням різних джерел подій, якщо їхні результати агрегуються та корелюються в SIEM. Як демонстраційний приклад у роботі розглянуто сценарій виявлення та реагування на SSH brute-force, оскільки він є наочним, відтворюваним і водночас практично значущим для сучасних інформаційних систем.

Метою статті є обґрунтування та експериментальна перевірка архітектурного підходу до автономного оркестрованого реагування на інциденти, у межах якого результати аналітики в

SIEM використовуються для формування інцидентного запису, запуску процедур реагування та журналювання їх виконання.

1. Теоретичні основи інтеграції SIEM та системи оркестрації

У сучасних інформаційних системах важливо не лише виявляти інциденти в SIEM, а й швидко та коректно перетворювати результат аналітики на практичну дію в інфраструктурі. Це особливо актуально в середовищах, де події надходять із різних джерел журналів операційних систем і сервісів, мережевих сенсорів, сканерів вразливостей та інших засобів моніторингу. Саме тому доцільним є використання проміжного архітектурного модуля, який поєднує аналітичний рівень SIEM із системою оркестрованого реагування та забезпечує керований перехід від результатів аналізу до виконання процедури на цільовому активі. [1,3,6]

Запропонований підхід поєднує аналітичний рівень SIEM, модуль-міст, систему оркестрації та контур керованої довіри до активів. Для цього інцидент подається у стандартизованому вигляді, повторні спрацювання контролюються через унікальний ключ інциденту та інтервал блокування повторного запуску, а безпечний доступ до кінцевих вузлів забезпечується без ручного підтвердження SSH-з'єднання під час інциденту. Загальну схему взаємодії цих компонентів наведено на рисунку 1.

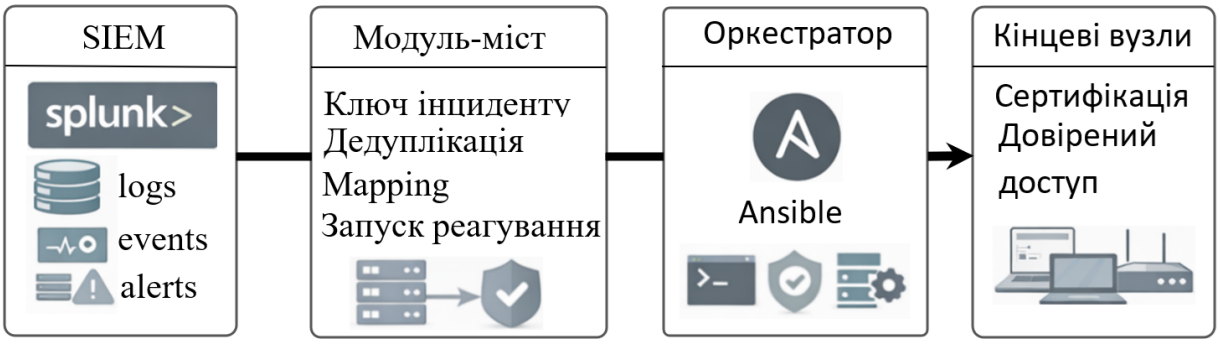


Рис. 1 Архітектура системи оркестрованого реагування на інциденти в SIEM-середовищі
Fig. 1 Architecture of an orchestrated incident response system in a SIEM environment

У межах запропонованої архітектури SIEM-шар виконує функції збору, індексації та аналітичної обробки подій, модуль-міст відповідає за інтерпретацію результатів пошуку, формування інцидентного запису, контроль повторних запусків і передачу параметрів до оркестратора, а система оркестрації виконує безпосередні дії реагування на визначених активах. Окремо функціонує керований контур довіри, побудований на основі SSH Host CA, який забезпечує безпечне підключення до кінцевих вузлів без ручного підтвердження ключів під час інциденту. Результати виконання процедур реагування фіксуються в журналі подій, що може виступати основою для подальшого аудиту та, за потреби, повторної індексації в SIEM, що дозволяє будувати точні умови запуску сцинаріїв реагувань.

Компонент SIEM і формування виходу аналітики. SIEM у запропонованому контурі виконує роль «точки нормалізації»: незалежно від того, чи подія походить із журналів ОС, мережевого сенсора або прикладного сервісу, вона потрапляє в єдиний аналітичний простір із можливістю кореляції. Детектор описується як запит до індексованих подій, а його результат має бути достатньо структурованим, щоб стати входом для реагування. Для Splunk типовим є опис детектора мовою SPL і виконання пошуку як у веб-інтерфейсі, так і через CLI або механізми alerting/scheduling. [6,7,12]

Для демонстрації працездатності підходу використано сценарій виявлення невдалих спроб SSH-автентифікації на цільовому вузлі. У межах експерименту аналітика будується на пошуку подій типу «Failed password» у журналах хоста, після чого виконується вилучення імені користувача та IP-адреси джерела, далі агрегація подій за полями host, src_ip та user. Такий підхід дає змогу виділити повторювані невдалі спроби входу, які можуть свідчити про brute-force активність. У запропонованому прикладі порогове спрацювання встановлено на рівні 5 невдалих спроб, що дозволяє виділити повторювану активність із ознаками brute-force та сформувати структурований результат для подальшого реагування. Отриманий результат далі

використовується для формування інцидентного запису та запуску відповідної процедури реагування. [6,7]

Лістинг 1. Приклад SPL-запиту
Listing 1. Example of an SPL query

```
index=* host=ub7* "Failed password"
| rex field=_raw "Failed password for (?:invalid user )?(?<user>\S+) from (?<src_ip>\d{1,3}(?:\.\d{1,3}){3})"
| stats count as failed_attempts latest(_time) as last_seen by host src_ip user
| where failed_attempts >= 5
| sort - failed_attempts
```

Наведений запит демонструє, як на основі журналів SSH можна виявити повторювані невдалі спроби автентифікації та звести їх до структурованого аналітичного результату. Агрегація за полями host, src_ip та user дає змогу отримати компактний набір ознак, достатній для формування інцидентного запису, оцінювання інтенсивності події та подальшого запуску процедури реагування. Використання порогової умови дозволяє відокремити поодинокі помилки входу від активності, що має ознаки brute-force, а сортування результатів спрощує пріоритезацію виявлених випадків. Таким чином, цей запит виступає не лише засобом виявлення, а й джерелом вхідних даних для модуля-міста, який перетворює результат SIEM-аналітики на керовану дію в інфраструктурі. Результат SPL-агрегації та поля, які формують інцидентний запис, наведено на рисунку 2 [5, 6].

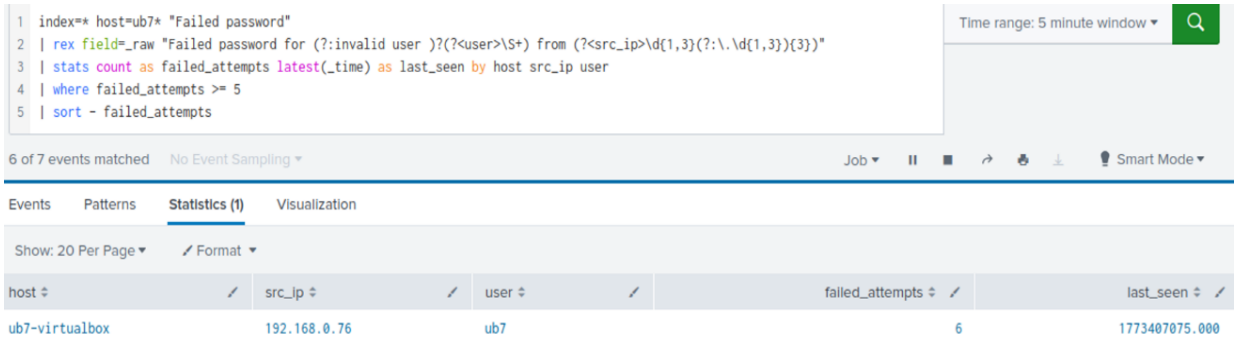


Рис. 2 Результат агрегації подій невдалих SSH-входів у Splunk
Fig. 2 Result of Aggregating Failed SSH Login Events in Splunk

Результат виконання запиту в SIEM містить поля, достатні для переходу від виявлення до реагування. Значення host описує цільовий актив, src_ip є адресою яка ініціювала подію, user є обліковим записом, щодо якого фіксуються спроби входу, failed_attempts це інтенсивність події, а last_seen її часову актуальність. На основі цього структурованого результату модуль-міст формує інцидентний запис, який використовується для логічної перевірки, оркестрації реагування та журналювання результатів. Саме в такому вигляді результат SIEM-аналітики переходить від етапу виявлення до етапу практичного реагування [6].

Таким чином можна сказати, що SIEM доцільно використовувати як централізований аналітичний рівень, у межах якого події з різнорідних джерел можуть бути зведені до узгодженого результату пошуку й кореляції. Для переходу від аналітики до практичної дії в інфраструктурі необхідний проміжний модуль, що забезпечує формалізацію інциденту, узгодження активів, керування повторними спрацюваннями та передавання параметрів до системи оркестрації. Така побудова створює основу для універсального контуру автономного реагування на інциденти, зафіксовані з використанням різних джерел подій [1, 2, 8].

Принципи формування інцидентного запису та функціонування модуля-міста

Робота модуля-міста визначається набором налаштувань, які задають правила інтерпретації аналітичного результату. До них належать назва правила виявлення, аналітичний запит, перелік потрібних полів, шаблон формування ключа інциденту, параметри контролю повторних запусків, спосіб узгодження активів із inventory та прив’язка до відповідної процедури реагування. Це

дозволяє модулю-місту працювати не як окремому скрипту під один сценарій, а як узагальненому механізму, який можна адаптувати до різних типів інцидентів шляхом зміни конфігурації. [5,6,7,8]

У межах реалізованого прототипу інцидентний запис формується у вигляді впорядкованого набору параметрів, що надалі передаються до модуля прийняття рішення та системи оркестрації. Параметри інцедентного запису отримуються зі SIEM системи, та формуються у json форматі після виявлення настання відповідного інциденту.

Лістинг 2. Приклад інцидентного запису у форматі JSON
Listing 2. Example of an Incident Record in JSON Format

```
{
  "rule_name": "ssh_bruteforce_multi_host",
  "host": "ub7-virtualbox",
  "src_ip": "192.168.0.76",
  "user": "ub7",
  "failed_attempts": 5,
  "last_seen": "1772905855.000"
}
```

Після формування інцидентного запису модуль-міст переходить до етапу логічної перевірки, у межах якого для кожного виявленого випадку формується унікальний ключ інциденту. У межах реалізованого прототипу такий ключ формується на основі назви правила та основних атрибутів події цільового активу, джерела активності й облікового запису користувача. Приклад формування такого ключа наведено в лістингу 3 [3, 5].

Лістинг 3. Приклад формування унікального ключа інциденту
Listing 3. Example of Generating a Unique Incident Key

```
incident_key = (
    f"ssh_bruteforce_multi_host|"
    f"host={host}|src_ip={src_ip}|user={user}"
)
```

Слід зазначити, що наведене формування унікального ключа інциденту (incident_key) не є жорстко прив’язаним лише до сценарію SSH brute-force. У межах повноцінної програмної реалізації генерація таких ключів має виконуватися автоматизовано на основі шаблону, що відповідає конкретному типу інциденту. Для різних класів подій склад інцидентного ключа може відрізнятися, оскільки він повинен включати саме ті атрибути, які дозволяють однозначно відокремити новий випадок від повторного спрацювання. Отже, зі збільшенням кількості підтримуваних сценаріїв реагування система має масштабувати і механізм формування ключів, зберігаючи при цьому єдиний принцип їх побудови [3].

Після формування ключа інциденту модуль виконує перевірку умов запуску. Якщо подія з таким самим ключем не оброблялася раніше або для неї вже завершився інтервал блокування повторного запуску, система автономно ініціює відповідну процедуру реагування. Якщо ж інцидент уже був оброблений у межах заданого інтервалу, повторний запуск не виконується. Таким чином, модуль-міст забезпечує автономне оркестроване реагування, за якого після виявлення інциденту система самостійно приймає рішення щодо запуску процедури відповідно до визначених політик. Це дає змогу поєднати аналітичний результат SIEM із керованим виконанням дій у цільовій інфраструктурі та забезпечити стійкість, повторюваність і практичну придатність усього контуру реагування [3].

Наступним етапом після перевірки умов запуску є узгодження цільового активу з даними інвентаризації системи оркестрації. Це необхідно тому, що ідентифікатор host, отриманий із SIEM, не завжди повністю збігається з тим, як відповідний вузол описаний у системі керування інфраструктурою. У SIEM актив може бути представлений у вигляді hostname, FQDN, IP-адреси або іншого позначення, сформованого джерелом журналювання, тоді як в inventory оркестратора той самий вузол може бути заданий під іншим ім’ям або через окремий параметр підключення. Саме тому перед запуском процедури реагування необхідно виконати етап зіставлення, у межах

якого аналітичний ідентифікатор активу приводиться у відповідність до запису, придатного для системи оркестрації [8].

Таке узгодження дає змогу уникнути ситуацій, коли інцидент коректно виявлено, але реакція не може бути виконана через невідповідність між позначенням вузла в SIEM та його описом в inventory. У межах запропонованого підходу цей механізм виступає важливою ланкою між етапом аналітики та етапом практичного реагування, оскільки саме він забезпечує коректну передачу інциденту на рівень виконання дій. У більш загальному вигляді такий підхід може бути реалізований через таблицю відповідностей, інвентаризаційний довідник або інший механізм зіставлення активів, що дає змогу масштабувати систему на різні типи вузлів і різні джерела подій.

Узгодження активу з inventory ще не гарантує можливості безпечного виконання реакційної процедури. Після того як модуль-міст визначив цільовий вузол, система повинна отримати автономний, але безпечний адміністративний доступ до нього. Якщо оркестратор під час інциденту залежить від ручного підтвердження SSH-підключення, підхід втрачає автономність. Якщо ж довіра до вузла формується через автоматичне прийняття першого ключа, це створює ризик підміни хоста. Саме тому безпечний доступ до кінцевих вузлів слід розглядати як одну з центральних умов працездатності всього контуру реагування [1,8,10].

Для розв’язання цієї задачі в архітектурі системи передбачено окремий контур довіри, який включає SIEM-сервер, сервер оркестрації, сервер сертифікації та кінцеві вузли. Особливу роль у цій схемі відіграє сервер оркестрації, оскільки саме він поєднує аналітичний рівень із практичним виконанням дій реагування. Водночас безпечний доступ до кінцевого вузла забезпечується не під час інциденту, а завдяки попередньо сформованому довіреному середовищу. Для цього використовується SSH Host CA, що дозволяє відмовитися від небезпечної довіри до першого ключа та зберегти обов’язкову перевірку SSH-ключа хоста перед підключенням [8,9,10].

Отже, сертифікація має передувати автономному реагуванню: вузли, на які потенційно можуть поширюватися реакційні дії, повинні бути заздалегідь підготовлені до безпечної взаємодії з оркестратором. Саме завдяки цьому контур довіри не прив’язується до одного сценарію, а стає універсальною основою для автономного реагування на інциденти різних типів [1,2].

Після завершення етапів інтерпретації інциденту, перевірки політик запуску, узгодження цільового активу з inventory та встановлення довіреного контуру доступу система переходить безпосередньо до запуску процедури реагування. На цьому етапі сервер оркестрації передає до playbook параметри, сформовані на основі інцидентного запису. До таких параметрів можуть належати ідентифікатор цільового вузла, тип або назва правила, IP-адреса джерела активності, ім’я користувача, кількість зафіксованих спроб, час останнього прояву події, а також службові атрибути, необхідні для журналювання та подальшого аналізу. У результаті playbook отримує не абстрактну команду на виконання, а структурований набір даних, який безпосередньо пов’язаний із конкретним інцидентом [8,9].

Подальше виконання playbook на кінцевому вузлі залежить від типу інциденту та обраної політики реагування. У загальному випадку така процедура може виконувати одну або кілька дій: фіксацію факту реагування, збір додаткових артефактів, зміну параметрів захисту, тимчасове обмеження доступу або інші керовані адміністративні операції. У межах реалізованого прототипу для підтвердження працездатності підходу playbook виконує контрольну дію на цільовому вузлі та залишає артефакт, який дозволяє однозначно перевірити факт виконання реакції. Такий підхід є принципово важливим, оскільки дає змогу не лише ініціювати процедуру реагування, а й підтвердити, що вона справді була виконана на визначеному активі [1,8,9,11].

Таким чином, система оркестрації виступає завершальною ланкою переходу від аналітичного результату SIEM до практичної дії в інфраструктурі. Якщо SIEM формує ознаки інциденту, а модуль-міст приймає рішення щодо доцільності запуску, то playbook реалізує безпосереднє застосування обраного сценарію до цільового вузла. Саме на цьому етапі автономне оркестроване реагування набуває завершеного вигляду, оскільки виявлена подія не лише фіксується, а й перетворюється на конкретну контрольовану дію з можливістю подальшого аудиту.

Програмна реалізація модуля-міста та практична перевірка запропонованого підходу

Для практичної перевірки запропонованого підходу на поточному етапі було реалізовано прототип модуля-міста у вигляді Python-скрипта. Його призначення полягає у підтвердженні технічної здійсненності методу та відтворенні основної логіки взаємодії між SIEM-аналітикою і системою оркестрації. У межах реалізованого прототипу скрипт приймає структурований

результат аналітичного запиту, формує інцидентний запис, будує унікальний ключ інциденту, перевіряє інтервал блокування повторного запуску, виконує зіставлення активу з inventory та передає підготовлений набір параметрів до playbook. Подальші лістинги демонструють реалізацію цих етапів у прототипному скрипті.

Наведений інцидентний запис у лістингу 2 показує, що кожне поле виконує окрему функцію в подальшій логіці обробки. Значення rule_name використовується для прив'язки запису до конкретного правила виявлення та відповідної процедури реагування. Поле host визначає вузол, до якого потенційно має бути застосована реакція, а src_ip характеризує джерело події. Поле user уточнює, щодо якого облікового запису зафіксовано підозрілу активність. Значення failed_attempts дозволяє оцінити інтенсивність події, а last_seen задає її часовий контекст. Саме на основі цих атрибутів надалі формується унікальний ключ інциденту, перевіряються політики повторного запуску та готуються параметри для передачі до playbook.

Після формування інцидентного запису скрипт переходить до побудови унікального ключа інциденту, який використовується для логічного відокремлення нового випадку від повторного спрацювання. Такий ключ формується на основі найбільш суттєвих атрибутів події, що дозволяють однозначно ідентифікувати конкретний інцидент у межах обраного сценарію. У розглянутому прикладі до складу ключа входять назва правила, цільовий актив, джерело активності та обліковий запис користувача. Використання саме цих полів дає змогу пов'язати між собою всі повторювані прояви одного й того самого випадку та, водночас, відокремити їх від інших подібних подій. Таким чином, incident_key виступає не лише технічним ідентифікатором, а й основою для подальшої перевірки політик повторного запуску, що є необхідною умовою автономного оркестрованого реагування.

Після формування ключа інциденту скрипт виконує перевірку умов запуску реагування. Для цього використовується локальний стан системи, що зберігається в таблиці trigger_state бази SQLite. На основі пари rule_name та incident_key визначається, чи оброблявся такий інцидент раніше, а також чи не перебуває він у межах заданого інтервалу блокування повторного запуску. Якщо різниця між поточним часом і часом останнього спрацювання менша за значення cooldown_sec, повторний запуск не виконується. Такий механізм дозволяє уникнути багаторазового ініціювання однієї й тієї самої процедури реагування, зменшує навантаження на систему оркестрації та забезпечує керованість автономного реагування.

Лістинг 4. Перевірка інтервалу блокування повторного запуску
Listing 4. Checking the restart lockout interval

```
    def is_in_cooldown(
        conn: sqlite3.Connection,
        rule_name: str,
        incident_key: str,
        cooldown_sec: int,
    ) -> bool:
        cur = conn.execute(
            "SELECT last_triggered_epoch FROM trigger_state WHERE rule_name=? AND
incident_key=?",
            (rule_name, incident_key),
        )
        r = cur.fetchone()
        if not r:
            return False
        last_ts = int(r[0])
        now = int(dt.datetime.utcnow().timestamp())
        return (now - last_ts) < int(cooldown_sec)
```

У наведеному фрагменті перевірка виконується шляхом звернення до локальної таблиці стану trigger_state, де для кожної пари rule_name та incident_key зберігається час останнього успішного спрацювання. Якщо відповідний запис відсутній, система вважає інцидент новим. Якщо ж запис існує, скрипт порівнює поточний час із моментом останнього запуску та визначає, чи не потрапляє подія в межі інтервалу блокування повторного запуску.

Якщо інцидент уже оброблявся в межах заданого інтервалу, система не виконує повторний запуск і фіксує стан cooldown_skip у журналі. Якщо ж інцидент є новим або інтервал блокування вже завершився, формується набір параметрів для передачі до playbook, та ініціюється процедура реагування. Таким чином, саме на цьому етапі модуль-міст реалізує автономне оркестроване реагування, оскільки після виявлення події система самостійно вирішує, чи потрібно запускати відповідну дію, в подальшому можна реалізувати умови виконання, коли інцедент потребує додаткового підтвердження не лише за фактором часу, а й за особливостями інших подій, що дозволяє реагувати на специфічну кореляцію деяких подій.

Лістинг 5. Умовний перехід до запуску реагування
Listing 5. Conditional transition to trigger response

```
    if is_in_cooldown(conn, rule_name, incident_key, int(rule["cooldown_sec"])):
        log_json_line(
            log_file,
            {
                "ts": utc_now_iso(),
                "rule": rule_name,
                "status": "cooldown_skip",
                "incident_key": incident_key,
                "splunk_host": splunk_host,
                "target_host": mapped_target,
            },
        )
    continue
```

Окремим етапом роботи скрипта є зіставлення цільового активу, отриманого з результату SIEM-аналітики, з його представленням у системі оркестрації. Це необхідно, оскільки значення host у журналі подій не завжди збігається з тим, як відповідний вузол описаний в inventory Ansible. У межах реалізованого прототипу таке зіставлення виконується через таблицю відповідностей host_map, що дозволяє перетворити аналітичний ідентифікатор активу у формат, придатний для подальшої роботи з inventory і запуску playbook.

Лістинг 6. Приклад зіставлення ідентифікатора активу з даними inventory
Listing 6. Example of mapping asset identifier to inventory data

```
    def map_target_host(cfg: dict, splunk_host_value: str) -> str:
        host_map = cfg.get("host_map", {}) or {}
    return host_map.get(splunk_host_value, splunk_host_value)
```

У наведеному фрагменті скрипт звертається до словника host_map, у якому зберігаються відповідності між позначеннями активів у SIEM та системі оркестрації. Якщо для отриманого значення host існує явне зіставлення, використовується перетворене значення; якщо ж такого запису немає, скрипт залишає початковий ідентифікатор без змін.

Після коректного зіставлення цільового активу з даними inventory система переходить до підготовки параметрів для запуску playbook. На цьому етапі сервер оркестрації передає до процедури реагування структурований набір даних, сформований на основі інцидентного запису та результатів зіставлення активу. У результаті playbook отримує не абстрактну команду, а пов’язаний із конкретним інцидентом контекст, необхідний для виконання дії на цільовому вузлі.

Лістинг 7. Формування параметрів для передавання до playbook
Listing 7. Preparation of Parameters for Passing to the Playbook

```
def build_extra_vars(cfg: dict, rule: dict, row: dict, incident_key: str) -> dict:
    extra_vars = {}

    for var_name, field_name in (rule.get("extra_vars_map", {}) or {}).items():
        extra_vars[var_name] = row.get(field_name, "")

    extra_vars.update(rule.get("extra_vars_static", {}) or {})

    if "target_host" in extra_vars:
        extra_vars["target_host"] = map_target_host(cfg, str(extra_vars["target_host"]))

    extra_vars["incident_key"] = incident_key
    extra_vars["rule_name"] = rule["name"]
  return extra_vars
```

У наведеному фрагменті скрипт формує словник extra_vars, у який включаються як значення, отримані з результату SIEM-пошуку, так і службові параметри, визначені конфігурацією правила. Окремо до цього набору додаються incident_key та назва правила, що дозволяє зберегти зв’язок між процедурою реагування і конкретним інцидентом.

Перед запуском дії реагування bridge-скрипт формує виклик ansible-playbook, у якому поєднуються шлях до цільового playbook, джерело інвентаря та набір параметрів інциденту, підготовлених на попередніх етапах обробки події. Змінні інциденту передаються у форматі JSON через механізм extra vars, що дає змогу динамічно підставляти до playbook значення цільового вузла. У лістингу 8 наведено фрагмент функції, яка реалізує запуск playbook та повертає результат виконання для подальшого аналізу й журналювання.

Лістинг 8. Запуск playbook із параметрами інциденту
Listing 8. Running the Playbook with Incident Parameters

```
def run_playbook(cfg: dict, rule: dict, extra_vars: dict) -> subprocess.CompletedProcess:
    ansible_cfg = cfg["ansible"]
    cmd = [
        "ansible-playbook",
        rule["playbook"],
        "-i",
        ansible_cfg["inventory"],
        "-e",
        json.dumps(extra_vars, ensure_ascii=False),
    ]
    timeout = int(ansible_cfg.get("playbook_timeout_sec", 300))
  return subprocess.run(cmd, capture_output=True, text=True, timeout=timeout)
```

Цей фрагмент показує, що запуск playbook виконується безпосередньо зі скрипта модуля-міста. До команди передаються шлях до playbook, інвентаризаційний файл і сформований набір параметрів extra_vars, що забезпечує прив’язку дій оркестратора до конкретного інциденту та цільового вузла.

Завершальним етапом роботи скрипта є журналювання результату виконання процедури реагування. Його призначення полягає в тому, щоб зафіксувати не лише факт запуску дії, а й її підсумковий стан. У межах реалізованого підходу журнал формується у структурованому вигляді та містить часову мітку, назву правила, ключ інциденту, статус виконання, цільовий актив і фрагменти службового виводу. Завдяки цьому він забезпечує простежуваність обробки інциденту та придатний для подальшого аудиту.

Важливо, що журнал охоплює не лише успішні випадки, а й інші стани роботи системи, зокрема відсутність збігів, блокування повторного запуску, помилки пошуку в SIEM і невдале виконання playbook. Таким чином, він відображає не лише результат реагування, а й логіку

прийняття рішення системою, що дозволяє використовувати його для технічної діагностики та, за потреби, повторної індексації в SIEM.

Лістинг 9. Формування журналу результатів виконання playbook
Listing 9. Generation of a Log of Playbook Execution Results

```
    payload = {
        "ts": utc_now_iso(),
        "rule": rule_name,
        "incident_key": incident_key,
        "status": "playbook_success" if result.returncode == 0 else "playbook_failed",
        "rc": result.returncode,
        "playbook": rule["playbook"],
        "target_host": extra_vars.get("target_host", ""),
        "splunk_host": splunk_host,
        "row": row,
        "stdout_tail": (result.stdout or "")[-1500:],
        "stderr_tail": (result.stderr or "")[-1500:],
    }
log_json_line(log_file, payload)
```

У наведеному фрагменті журнал формується як структурований запис, що містить часову мітку, назву правила, ключ інциденту, статус виконання, код завершення, ідентифікатор playbook, цільовий актив та фрагменти службового виводу. Поле status дозволяє одразу відрізнити успішне виконання від невдалого, а збереження stdout_tail і stderr_tail спрощує подальший аналіз причин помилок без необхідності повторного відтворення інциденту. Така структура журналу робить його придатним як для локального технічного контролю, так і для подальшої інтеграції в засоби централізованого моніторингу.

Окремою перевагою є те, що журналювання безпосередньо пов’язується з інцидентним записом через incident_key. Це означає, що кожен запис про виконання або пропуск процедури реагування може бути однозначно співвіднесений із конкретним виявленим випадком. У практичному сенсі така прив’язка дає змогу аналізувати не лише окремі події, а й повний життєвий цикл інциденту: від моменту його виявлення в SIEM до завершення реакційної процедури або відмови від її запуску за визначеними політиками. Саме це і забезпечує доказовий характер журналу реагування.

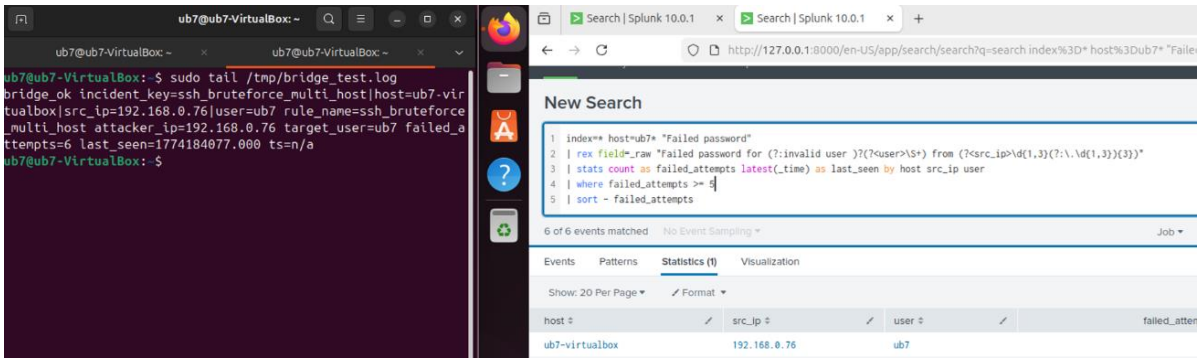


Рис. 3 Підтвердження виявлення інциденту у SIEM та виконання реакції на цільовому вузлі
Fig. 3. Confirmation of Incident Detection in the SIEM and Execution of the Response on the Target Host

На рисунку 3 наведено приклад виявлення повторюваних невдалих SSH-автентифікацій у Splunk, а також підтвердження того, що після цього на цільовому вузлі було виконано реакційну процедуру, результат якої зафіксовано у файлі bridge_test.log.

Отже, отримані результати підтверджують, що поєднання SIEM-аналітики з проміжним модулем-містом і системою оркестрації дає змогу реалізувати керований контур автономного реагування на інциденти. У ході дослідження показано, що результат аналітики в SIEM може бути послідовно перетворений на інцидентний запис, використаний для формування унікального ключа інциденту, перевірки політик повторного запуску, узгодження цільового активу з inventory та подальшого запуску сценарію реагування (playbook). Практична перевірка на прикладі SSH

brute-force підтвердила технічну здійсненність такого підходу: було реалізовано повний цикл від виявлення події в SIEM до виконання сценарію реагування на цільовому вузлі та фіксації результату в журналі реагування. Окремо підтверджено, що використання SSH Host CA дає змогу забезпечити безпечний автономний доступ до кінцевих вузлів без ручного підтвердження під час інциденту.

Висновки

По-перше, SIEM доцільно використовувати як централізований аналітичний рівень, у межах якого події з різнорідних джерел можуть бути зведені до узгодженого результату пошуку й кореляції. Для переходу від аналітики до практичної дії в інфраструктурі необхідний проміжний модуль, що забезпечує формалізацію інциденту, узгодження активів, керування повторними спрацюваннями та передавання параметрів до системи оркестрації. Саме така побудова створює основу для універсального контуру автономного реагування на інциденти, зафіксовані з використанням різних джерел подій. По-друге, система оркестрації є завершальною ланкою переходу від аналітичного результату SIEM до практичної дії в інфраструктурі. Якщо SIEM формує ознаки інциденту, а модуль-міст визначає доцільність запуску, то playbook реалізує безпосереднє застосування обраного сценарію до цільового вузла. Саме на цьому етапі автономне оркестроване реагування набуває завершеного вигляду, оскільки виявлена подія не лише фіксується і класифікується, а й перетворюється на конкретну контрольовану дію з можливістю подальшого аудиту її результатів. По-третє, поєднання SIEM-аналітики з проміжним модулем-містом і системою оркестрації забезпечує керований контур автономного реагування на інциденти. Результат аналітики в SIEM послідовно переходить в інцидентний запис, використовується для формування унікального ключа інциденту, перевірки політик повторного запуску, узгодження цільового активу з inventory та запуску сценарію реагування. Практична перевірка на прикладі SSH brute-force підтвердила технічну здійсненність такого підходу: реалізовано повний цикл від виявлення події в SIEM до виконання сценарію реагування на цільовому вузлі та фіксації результату в журналі реагування. Використання SSH Host CA забезпечує безпечний автономний доступ до кінцевих вузлів без ручного підтвердження під час інциденту.

Запропонований підхід вирізняється тим, що поєднує в єдиному контурі аналітичний рівень SIEM, модуль-міст, механізм контролю повторних запусків, засоби узгодження активів та систему оркестрації реагування. На відміну від рішень, орієнтованих на окремі типи телеметрії або вузькоспеціалізовані засоби захисту, така архітектура дозволяє будувати автономне реагування на інциденти, зафіксовані з використанням різних джерел подій, якщо їх результати агрегуються та корелюються в SIEM. У цьому контексті важливим є не лише саме виявлення події, а й перетворення результату аналітики на стандартизований інцидентний запис, придатний для подальшої логічної обробки, запуску процедури реагування та фіксації її результату.

Одержані результати розширюють уявлення про перехід від етапу виявлення інциденту до етапу практичного реагування в розподіленій інформаційній інфраструктурі. Запропонований підхід дає підстави розглядати автономне реагування не як набір окремих скриптів або локальних інтеграцій, а як цілісний керований контур із визначеними вхідними даними, умовами запуску, політиками повторного виконання та засобами аудиту. Практична цінність роботи полягає в тому, що така архітектура може бути використана як основа для реальних систем реагування, у яких необхідно не лише швидко перейти від виявлення до дії, а й зберегти керованість, відтворюваність і доказовість виконаних процедур.

Подальший розвиток роботи доцільно пов’язати зі створенням повноцінної програмної реалізації модуля-міста, розширенням бібліотеки сценаріїв реагування, підтримкою інших типів інцидентів, інтеграцією з довідниками активів та механізмами автоматичного збагачення інцидентного запису контекстною інформацією. Перспективним напрямом також є повторне індексування журналу реагування в SIEM для подальшого оцінювання ефективності політик і вдосконалення всього контуру автономного оркестрованого реагування.

СПИСОК ЛІТЕРАТУРИ

1. C. Pascoe, S. Quinn, K. Scarfone. The NIST Cybersecurity Framework (CSF) 2.0. NIST Cybersecurity White Paper 29. Gaithersburg, MD, USA : National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.CSWP.29. https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf
2. K. Rigopoulos, S. Quinn, C. Pascoe, J. Marron, A. Mahn, D. Topper. NIST Cybersecurity Framework 2.0: Resource & Overview Guide. NIST Special Publication 1299. Gaithersburg, MD, USA : National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.SP.1299. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1299.pdf
3. A. Nelson, S. Rekhi, M. Souppaya, K. Scarfone. Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile. NIST Special Publication 800-61 Rev. 3. Gaithersburg, MD, USA : National Institute of Standards and Technology, 2025. DOI: 10.6028/NIST.SP.800-61r3. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r3.pdf
4. European Union Agency for Cybersecurity (ENISA). ENISA Threat Landscape 2024. 2024. (дата звернення: 23.03.2026) URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024
5. MITRE ATT&CK. Brute Force, Technique T1110 - Enterprise. (дата звернення: 23.03.2026) URL: https://attack.mitre.org/techniques/T1110/
6. Splunk. Search | Splunk Enterprise. (дата звернення: 23.03.2026) URL: https://docs.splunk.com/Documentation/Splunk/9.4.2/SearchReference/Search
7. Splunk. savedsearches.conf | Platform. (дата звернення: 23.03.2026) URL: https://docs.splunk.com/Documentation/Splunk/9.4.2/Admin/Savedsearchesconf
8. Ansible Project. How to build your inventory. https://docs.ansible.com/projects/ansible/latest/inventory_guide/intro_inventory.html
9. Ansible Project. ansible-playbook. (дата звернення: 23.03.2026) URL: https://docs.ansible.com/projects/ansible/latest/cli/ansible-playbook.html
10. OpenBSD. ssh-keygen(1). (дата звернення: 23.03.2026) URL: https://man.openbsd.org/OpenBSD-7.6/ssh-keygen.1
11. Anderson R. Security Engineering: A Guide to Building Dependable Distributed Systems. 3rd ed. Hoboken, NJ, USA : Wiley, 2020. https://www.wiley.com/en-cn/Security+Engineering%3A+A+Guide+to+Building+Dependable+Distributed+Systems%2C+3rd+Edition-p-9781119642787
12. Bejtlich R. The Tao of Network Security Monitoring: Beyond Intrusion Detection. Boston, MA, USA : Addison-Wesley, 2004. https://www.informit.com/store/tao-of-network-security-monitoring-beyond-intrusion-9780321246776

Bratko Dmytro Viacheslavovych *undergraduate, cadet*
Institute of Special Communication and Information Protection of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 4 Verkhnokliuchova St., Kyiv, Ukraine, 03056
e-mail: loading.2285@gmail.com;
https://orcid.org/0009-0001-0863-9285

Kubrak Volodymyr Oleksandrovych *PhD student, Senior Lecturer of Special Department № 1*
Institute of Special Communication and Information Protection of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 4 Verkhnokliuchova St., Kyiv, Ukraine, 03056
e-mail: v.kubrak@kpi.ua;
https://orcid.org/0000-0001-8877-5289

Matiiko Aleksandra Andriivna *PhD student, Associate Professor of Special Department № 1*
Institute of Special Communication and Information Protection of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 4 Verkhnokliuchova St., Kyiv, Ukraine, 03056
e-mail: alexm1710@ukr.net;
https://orcid.org/0000-0002-6947-5958

Autonomous orchestrated incident response system based on SIEM

Relevance. Modern information systems generate security events from various sources, including operating system and service logs, network sensors, vulnerability scanners, and other monitoring tools. In such conditions, SIEM enables the centralized collection, indexing, and correlation of telemetry; however, the transition from analytical results to practical response often remains insufficiently formalized. This leads to delays, dependence on manual actions, difficulties in ensuring the repeatability of procedures, and the absence of a unified mechanism for confirming executed response actions. An additional challenge is the provision of secure autonomous access to endpoints during an incident, when both manual confirmation of SSH connections and insecure trust on first use are unacceptable. In this context, the development of an architectural bridge between SIEM analytics and an orchestration system is highly relevant, as it can ensure controlled, repeatable, and auditable incident response regardless of the original source of events.
Goal. The purpose of this work is to substantiate and experimentally validate an architectural approach to autonomous orchestrated incident response, in which the results of SIEM analytics are transformed into a structured incident record and then used to initiate response procedures in an orchestration system. To achieve this goal, the detector is described in a declarative form, the incident record is standardized, repeated triggering is controlled through a unique incident key and a re-execution lockout interval, target assets are aligned with inventory data, execution results are logged, and secure access to endpoints based on SSH Host CA is implemented. The scenario of detecting and responding to an SSH brute-force attack was chosen as a demonstration case.
Results. As a result of the study, an architectural approach that combines SIEM analytics with the automated execution of response actions on target assets was developed and experimentally validated. It was shown that the result of an analytical query in a SIEM can be consistently transformed into an incident record, used to construct a unique incident key, verify re-execution policies, align the target asset with the inventory, and transfer parameters to a playbook. The implemented prototype confirmed the technical feasibility of building a complete cycle from event detection in the SIEM to the execution of a response procedure on an endpoint and the recording of the result in a structured log. It was also confirmed that the use of SSH Host CA makes it possible to provide secure autonomous access to endpoints without manual confirmation during an incident. The obtained results further demonstrated that the proposed architecture can be scaled to other response scenarios provided that the detection rules and execution procedures are adapted accordingly.
Conclusions. The obtained results confirm that the integration of SIEM analytics with an orchestration system makes it possible to implement a controlled framework for autonomous incident response. The result of SIEM analytics is transformed into an incident record, which is then used to control repeated executions, align the target asset with the inventory, and initiate a response scenario. Practical validation based on the SSH brute-force case confirmed the technical feasibility of this approach: a complete cycle was implemented, from event detection to response execution and the recording of its result in the log. The proposed architecture is suitable for responding to incidents identified from various event sources, provided that their results are aggregated and correlated in the SIEM. The use of SSH Host CA ensures secure autonomous access to endpoints without manual confirmation during an incident. Further development of this work should be associated with the software implementation of the bridge module, the expansion of the response scenario library, and the transfer of response logs back to the SIEM for further analysis.
Keywords: *Splunk, Ansible, SIEM, orchestrated response, autonomous response, SSH brute-force, SSH Host CA.*

Надійшла у першій редакції 07.02.2026; рекомендована до друку 11.03.2026; опублікована 03.04.2026.
The first version has been received on 07.02.2026; was recommended for printing on 11.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-03

UDC 004.85:510.6:004.421.2

Yevdokymov Oleksandr *PhD Student, Department of Mathematical Modeling and Artificial Intelligence, National Aerospace University "Kharkiv Aviation Institute", Vadym Manko St., 17, Kharkiv, Ukraine, 61070*
e-mail: o.yevdokymov@khai.edu
https://orcid.org/0009-0008-9687-6344

Luchsheva Oksana *Senior Lecturer, Department of Software Engineering National Aerospace University "Kharkiv Aviation Institute", Vadym Manko St., 17, Kharkiv, Ukraine, 61070*
e-mail: o.luchsheva@khai.edu ;
https://orcid.org/0000-0003-3855-2815

A Mathematical Model of Automatic Verification of Formalized Proofs and a Conservative Presentation Interface over Lean

Relevance. Interactive theorem provers such as Lean have fundamentally transformed the verification of mathematical statements by transferring the final control of correctness from the human mathematician to a small trusted kernel. Nevertheless, their widespread adoption in research and higher education is still limited by a significant gap between the strict formal language of the kernel, based on dependent type theory, and the usual intuitive mathematical notation, symbols, and natural-language reasoning used by mathematicians and students. This discrepancy creates serious obstacles for teaching proof writing, formalizing new results, and developing effective intelligent educational systems.
Objective. The purpose of this paper is to construct a rigorous mathematical model of automatic verification of formalized proofs in Lean and to provide a formal justification for building a conservative human-oriented presentation interface together with an untrusted generative pedagogical component over the trusted kernel without any loss of mathematical rigor.
Methods. The study relies on dependent type theory, metatheoretic properties of the calculus of constructions (soundness, strong normalization, decidability of type checking), models of elaboration and backward projection between the presentation language and the kernel, and formal predicates of term-typing correctness. Interface conservativity is expressed as the inclusion of the set of derivable statements of the presentation layer in the set of kernel theorems. The safety of the generative AI-component is defined by the condition that every executable artefact must be projected back to the kernel and re-verified by the trusted type-checker.
Results. The developed model defines the formalizable fragment of mathematical discourse $\mathcal{ML}$, the kernel-verification predicate VerL, the course formalizability functional $\mathbb{F}(\mathcal{S})$, a complete conservative interface model, and a safety condition for the generative extension. It is proved that syntactic sugar, macros, notation abbreviations, and AI-generated educational artefacts do not enlarge the set of derivable statements as long as the trusted Lean kernel remains unchanged and all executable fragments are re-checked by the kernel. An illustrative example with four theorems from algebra and group theory (uniqueness of the identity element, intersection of subgroups, kernel of a homomorphism is normal, and binomial identity) demonstrates that human-oriented notation is fully elaborated into kernel terms while preserving mathematical meaning and verifiability. The interface fidelity reaches the maximum value of 1.
Conclusions. The obtained results demonstrate that the formal language of Lean can be equipped with a convenient human-oriented interface and intelligent educational services suitable for scientific and instructional use without compromising rigor. The only criterion of truth remains repeated verification by the small trusted kernel. The proposed separation between verified kernel knowledge, conservative presentation layer, and untrusted generative services creates a reliable foundation for building intelligent educational systems in higher mathematics, discrete mathematics, algebra, and mathematical logic, where formal correctness is guaranteed independently of the quality of generated explanations and exercises.
Keywords: *Lean, dependent types, automatic proof verification, formalized mathematics, conservative interface, generative pedagogical module.*

How to quote: O. Yevdokymov, and O. Luchsheva, "A Mathematical Model of Automatic Verification of Formalized Proofs and a Conservative Presentation Interface over Lean", *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 69, pp. 33–40, 2026. https://doi.org/10.26565/2304-6201-2026-69-03

Як цитувати: Yevdokymov O., and Luchsheva O. A Mathematical Model of Automatic Verification of Formalized Proofs and a Conservative Presentation Interface over Lean. *Вісник Харківського національного університету імені В. Н. Каразіна, серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління».* 2026. вип. 69. С. 33–40. https://doi.org/10.26565/2304-6201-2026-69-03

1. Introduction

Modern interactive theorem provers have transformed the machine verification of mathematical statements from a purely logical task into a practical instrument for the formalization of substantial fragments of contemporary mathematics. Lean is characterized by a small trusted kernel, dependent types as the basis of its internal language, and an advanced mechanism for syntax extension, elaboration, tactics, and inference support, while the mathlib library accumulates a large corpus of already verified definitions, lemmas, and theorems [1–3]. The theoretical foundations of such systems are provided by intuitionistic type theory, the calculus of constructions, and the correspondence "propositions as types, proofs as terms" [4–6].

For research and education this means that the control of mathematical correctness can be transferred from the human agent to the system kernel, whereas the human remains responsible for formulating ideas, choosing a strategy, selecting notation, and constructing an explanatory layer. Experience with Lean in proof education, as well as work on controlled natural language and human-oriented interfaces such as Verbose Lean and Waterproof, shows that the principal obstacle is not verification itself but rather the mode of interaction between the mathematician and the formal language [7–12].

At the same time, practical intelligent educational systems impose two nontrivial requirements. First, the mathematically trusted layer must be strictly separated from untrusted service extensions. Second, it must be formally justified why ordinary mathematical notation, macros, syntactic sugar, or AI-generated explanations do not alter the set of derivable theorems whenever final verification is performed by the kernel. Existing studies have already taken steps toward the operationalization of the formalizability of mathematical tasks for intelligent tutoring systems, as well as toward the representation of mathematical expressions by structural trees [13–15]; however, the mathematical model of a conservative human-oriented interface over Lean still requires a separate justification.

The aim of the paper is to construct a mathematical model of the automatic verification of formalized proofs in Lean and to rigorously justify the conservativity of a human-oriented presentation interface and an untrusted generative component with respect to the trusted kernel of the system.

2. Formal problem statement

Let L be a Lean-like formal system built over dependent type theory. At the minimal level of abstraction it is convenient to represent it by the tuple

$$L = (\Sigma, Ctx, Term, Type, \vdash, \equiv, K) \tag{2.1}$$

where Σ is the signature of constants and inductive objects, *Ctx* is the set of contexts, *Term* is the set of terms, *Type* is the set of types, $\vdash$ is the typing relation, $\equiv$ is definitional equality, and *K* is the trusted verification kernel. The basic judgment has the form

$$\Gamma \vdash t : T \tag{2.2}$$

where Γ is the context of local hypotheses and instances, *t* is a term, and *T* is a type. By the Curry-Howard correspondence, a proof of a statement T is represented by a term p of that type [4]–[6].

$$\mathrm{Th}(L) = \{ T \in \mathrm{Type} \mid \exists p \in \mathrm{Term}: \varnothing \vdash p : T \} \tag{2.3}$$

Let $\mathcal{M}$ denote the set of statements formulated in ordinary mathematical language, and let $\mathcal{M}_L \subset \mathcal{M}$ be the fragment for which an adequate encoding into L exists. Then

$$\mathcal{M}_L = \{\sigma \in \mathcal{M} \mid \exists \varphi(\sigma) \in \mathrm{Type},\ \exists p \in \mathrm{Term}: \varnothing \vdash p : \varphi(\sigma)\} \tag{2.4}$$

Here φ is the mapping of a mathematical statement into a formal type. For a finite set of statements $\mathcal{S} = \{\sigma 1, \ldots, \sigma N\}$, we introduce the formalizability coefficient

$$\mathfrak{F}(\mathcal{S}) = \frac{1}{|\mathcal{S}|} \sum_{\sigma \in \mathcal{S}} \mathbf{1}_{\sigma \in \mathcal{M}_L} \tag{2.5}$$

The quantity $\mathfrak{F}(\mathcal{S})$ generalizes the intuitive claim about what proportion of a course, topic, or problem bank admits full formalization and automatic verification [13]. In order to avoid an excessively strong metaphysical thesis about the formalization of "any" proof without qualifications, universality of automatic verification will henceforth mean universality on the set $\mathcal{M}_L$ only.

A formalization will be considered adequate if it preserves the mathematical meaning of the statement:

$$[\![\sigma]\!]_{\mathcal{M}} = [\![\varphi(\sigma)]\!]_L \tag{2.6}$$

where $[\![\cdot]\!]_{\mathcal{M}}$ and $[\![\cdot]\!]_L$ denote the semantic interpretations at the levels of informal and formal mathematics, respectively. Equality (1.6) is the condition of faithfulness of translation from the language of the mathematician into the language of the kernel.

3. Model of automatic proof verification

At the surface level, the user may build a proof by a tactic script, a term expression, or a combination of both. After elaboration, the surface text is reduced to a kernel term. Therefore, the model of automatic verification can be defined through the Boolean predicate

$$\mathrm{Ver}_L(T, p) = 1 \Leftrightarrow K \text{ accepts } p{:}\, T \tag{3.1}$$

In the opposite case, we assume that *VerL(T, p) = 0*. From the mathematical point of view, any formal proof may be represented as a derivation tree

$$D = (V_D, E_D, r, \lambda), \qquad \lambda(v) = \Gamma_v \vdash t_v{:}\, T_v \tag{3.2}$$

where the vertices of the tree correspond to local judgments and the edges correspond to applications of inference rules. If the leaves of *D* are axioms or directly checkable constructor steps, and the root coincides with $\vdash p : T$, then *D* encodes a completed proof of the theorem *T*.

Proposition 1. For any closed $T \in Type$ and $p \in Term$, the predicate *VerL(T, p)* is computable.

Proof. Within the kernel, correctness checking reduces to solving the typing problem $p : T$ while taking definitional equality into account. For the calculus of constructions and its practical restrictions used in Lean, the type-checking procedure is algorithmic; therefore, there exists an algorithm that terminates on every closed input with value 0 or 1 [1], [2], [4], [5].

Theorem 1. Let $\sigma \in \mathcal{M}L$ and suppose the adequacy condition (1.6) holds. If $VerL(\varphi(\sigma), p) = 1$, then the statement σ is *true* in the intended mathematical interpretation.

Proof. From $VerL(\varphi(\sigma), p) = 1$ we obtain $\vdash p : \varphi(\sigma)$. By the metatheoretic soundness of the kernel, this implies the semantic truth of the formal statement $\varphi(\sigma)$ in every model of the system *L* compatible with the context. By the faithfulness of translation (2.6), the truth of $\varphi(\sigma)$ is equivalent to the truth of σ in the mathematical interpretation.

Thus, any mathematical proof belonging to the formalizable fragment $\mathcal{M}L$ and equipped with a constructed proof object *p* is automatically checkable by the kernel. This is the mathematical meaning of the thesis of automatic proof verification in Lean.

4. Conservative model of a human-oriented presentation interface

To make work with a formal system acceptable for a researcher, teacher, or student, a presentation layer *P* is introduced over the kernel language: ordinary mathematical notation, abbreviated forms of quantifiers, symbols such as $x \in H$, $\ker f$, $H \trianglelefteq G$, templates of natural-language steps, and macros. The properties of Lean 4 as an extensible environment with a programmable parser, elaborator, and pretty-printer make such an extension natural [2], while educational systems such as Verbose Lean and Waterproof demonstrate the pedagogical value of human-oriented syntactic layers [11], [12].

Let the interface be specified by the tuple

$$I = (P, C, E, R, \Pi, K) \tag{4.1}$$

where *P* is the presentation language, *C* is the Lean kernel language, $E : P \rightharpoonup C$ is elaboration, $R : C \to P$ is rendering into a readable form, $\Pi : P \rightharpoonup C$ is the projection of executable fragments into the kernel, and *K* is the trusted kernel. The set of statements accessible to the user in language *P* is defined as

$$\mathrm{Th}(P) = \{\, s \in P \mid \exists p \in \mathrm{Term}{:}\, \mathrm{Ver}_L(E(s), p) = 1 \,\} \tag{4.2}$$

Then the formal expression of interface conservativity is the inclusion

$$E\big(\mathrm{Th}(P)\big) = \mathrm{Th}(C) \cap \mathrm{Im}(E) \tag{4.3}$$

Theorem 2. Suppose every extension of the syntax of *P* is eliminated by the function *E* into terms of the language *C*, while the rules of the kernel *K* remain unchanged. Then *P* is a conservative extension over *C*.

Proof. Let $s \in P$ be derivable. By definition (4.2), there exists p such that $VerL(E(s), p) = 1$, that is, *E(s)* belongs to *Th(C)*. Hence no statement becomes derivable solely due to the syntactic extension.

Conversely, any term $c \in C$ may be regarded as a trivial element of the presentation language or may be rendered into it via *R*. Therefore, *P* does not enlarge the set of theorems but only changes their mode of expression.

For practice, the stability of the cycle "kernel → readable form → kernel" is also useful:

$$\mathrm{NF}\left(E\left(R(c)\right)\right) = \mathrm{NF}(c), \qquad c \in C \tag{4.4}$$

where *NF* denotes normalization up to *α*-, *β*-, *δ*-, *ι*-, and *ζ*-conversions. If equality (3.4) holds on the relevant class of terms, the interface preserves not only the set of derivable statements but also syntactic identity after normalization. For a finite test set $X = \{c1, \ldots, cm\}$, interface fidelity may be measured by the indicator

$$Q_I(X) = \frac{1}{|X|}\sum_{c\in X} \mathbf{1}_{\mathrm{NF}\left(E\left(R(c)\right)\right)=\mathrm{NF}(c)} \tag{4.5}$$

In the case $QI(X) = 1$, the interface is fully faithful on the chosen set of formal objects. Several typical examples of translating ordinary mathematical notation into kernel constructions are given in Table 1.

Табл. 1. Відображення звичайної математичної нотації в конструкції Lean-ядра
Table 1. Mapping ordinary mathematical notation to Lean kernel constructions

Mathematician's language	Kernel form / elaboration	Purpose
∀ x ∈ A, P(x)	∀ x, x ∈ A → P x	Bounded quantifier abbreviation
x ∈ ker f	f x = 1	Expansion of the definition of the kernel of a homomorphism
H ⊴ G	Normal H	Abbreviated notation for a normal subgroup
a + b = b + a	Eq (a + b) (b + a)	Equality as a type object
"Let x be arbitrary …"	fun x => ...	Natural-language introduction of a quantifier

The general architecture of the proposed conservative interface is shown in Fig. 1

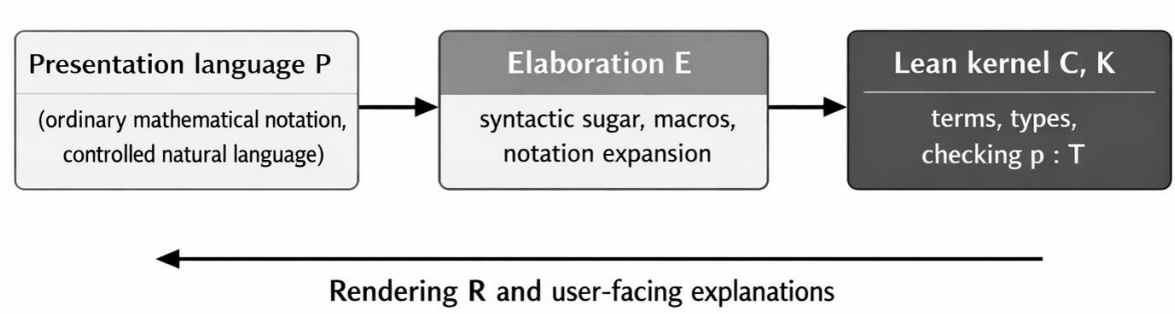


Рис. 1. Консервативна модель інтерфейсу над довіреним ядром
Fig. 1. Conservative interface model over the trusted kernel

5. Generative component as an untrusted pedagogical extension

The construction of a convenient mathematical interface is compatible with the use of AI provided that AI does not enter the trusted kernel. Let *A* be a nondeterministic operator that receives a kernel-level formal object as input and returns a set of pedagogical artefacts: explanations, examples, micro-lectures, generated exercises, and occasionally fragments of executable Lean code. Formally, let

$$A: C \rightsquigarrow Y, \qquad Y = Y_{\mathrm{text}} \cup Y_{\mathrm{exec}} \tag{5.1}$$

The elements of the set Y_{text} are purely textual and therefore do not directly affect mathematical correctness. By contrast, the elements of Y_{exec} may contain Lean fragments or other executable objects; therefore, they must be returned to the kernel via the projection function *Π*. We call the extension *A* safe if the condition

$$\mathrm{Safe}(A) \Leftrightarrow \forall c{:}T\ \forall y \in A(c){:}\ y \in Y_{\mathrm{exec}} \Rightarrow \exists c' = \Pi(y)\ \wedge\ \mathrm{Ver}_L(T, c') = 1 \tag{5.2}$$

Theorem 3. If the AI-extension *A* is safe in the sense of (5.2), then adding it to the interface does not change the set of derivable theorems and does not weaken the soundness of the system.

Proof. Artefacts from Y_{text} are non-executable and therefore logically inert: they do not generate new objects of type theorem/proof. For every $y \in Y_{exec}$, condition (5.2) requires the existence of $c' = \Pi(y)$, which is accepted by the kernel. Hence all executable AI outputs pass through the same verification channel as ordinary user proofs. It follows that the set of accepted statements coincides with $Th(C) \cap Im(E)$, whereas soundness is determined solely by the kernel *K*.

The safe interaction between the generative component and the kernel is illustrated in Fig. 2.

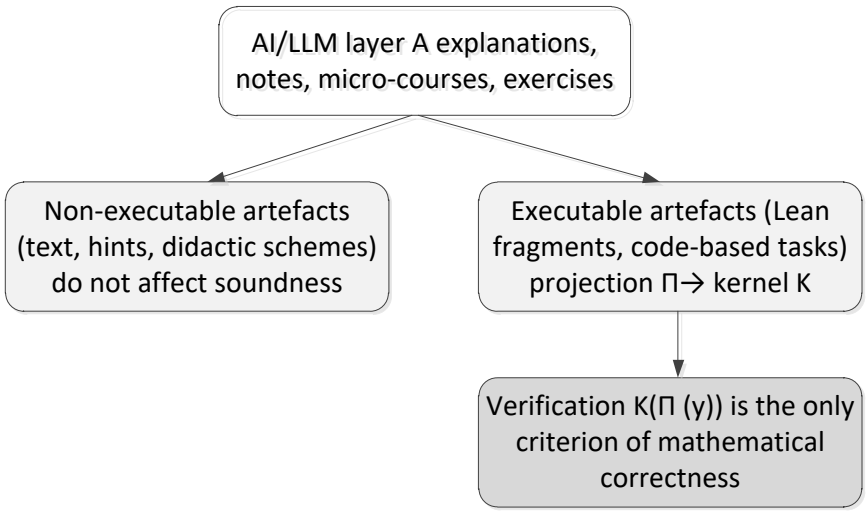


Рис. 2. Безпечне генеративне розширення: лише виконувані артефакти, що проектуються назад у ядро, можуть впливати на коректність
Fig. 2. Safe generative extension: only executable artefacts projected back into the kernel may affect correctness

It is precisely this property that provides a mathematical justification for an AI-generated learning module built over formally verified theorems. The generative component may produce explanations, short notes, exercises, and examples; however, every executable fragment must be returned to the kernel via *Π* and verified by the predicate $K(\Pi(y))$. Therefore, mathematical rigor is determined not by the stylistic quality of generation but by the architecture of interaction with the trusted kernel.

$$\mathrm{Course}_A(\Theta) = \bigcup_{T\in\Theta} \{A(T),\, R(T),\, \mathrm{examples}(T),\, \mathrm{tasks}(T)\} \tag{5.3}$$

In other words, the rigor of the course is determined not by the textual quality of the AI output but by the fact that all executable artefacts from the sets *tasks(T)* or *examples(T)* are returned to the kernel through *Π* and checked by it. In this architecture, an LLM is not a proven "mathematician" in the trusted sense; it is merely a generator of human-oriented presentation, whereas the final judge of correctness is the small trusted kernel [1], [2].

6. Illustrative example and didactic implications of formalization

To illustrate the model, consider a small set of theorems $\Theta = \{T1, T2, T3, T4\}$, covering both elementary algebra and basic group theory. All statements were selected so that they admit natural representation both in ordinary mathematical language and in compact Lean-like notation. Let us introduce the proportion of automatically verified statements in the set Θ:

$$\eta(\Theta) = \frac{1}{|\Theta|}\sum_{T\in\Theta} \mathbf{1}_{\exists p:\mathrm{Ver}_L(T,p)=1} \tag{6.1}$$

If formal proof objects have been constructed for all theorems in the set, then $\eta(\Theta) = 1$. In this illustrative case, precisely such a situation is considered: the purpose is not empirical measurement but a demonstration of how a human-oriented interface and kernel-level verification interact on specific mathematical examples.

For the same set one may evaluate interface fidelity using functional (4.5). If presentation macros for the symbols $\in$, *ker*, $\trianglelefteq$, as well as abbreviations of the form "$\forall x \in A$", are translated by the elaborator into standard Lean terms without change of meaning, then for the selected set $X = \{T1, T2, T3, T4\}$ we have

QI(X) = 1. Hence the user operates with a familiar language, while the kernel operates with the canonical form.

The four selected theorems together with their formal representations and the role of the interface are shown in Table 2.

Табл. 2. Ілюстративний набір формалізованих тверджень
Table 2. Illustrative set of formalized statements

Label	Mathematical statement	Formal image type	VerL	Role of the interface
T1	In a group, the identity element is unique	∀ e_1 e_2 : G, Id e_1 → Id e_2 → e_1 = e_2	1	Abbreviated notation Id, G
T2	The intersection of two subgroups is a subgroup	Sub(H) → Sub(K) → Sub(H ∩ K)	1	Notation ∩ and x ∈ H
T3	The kernel of a homomorphism is a normal subgroup	Hom(f : G → G′) → Normal(ker f)	1	Abbreviations ker, ⊴
T4	Binomial identity	∀ x y : R, (x+y)^2 = x^2 + 2xy + y^2	1	Standard algebraic notation

It is characteristic that the same principle extends to the construction of a learning module. For example, for *T3* the generative component may produce a short explanation such as “to prove that the kernel is normal, it is sufficient to show closure under conjugation”, as well as auxiliary exercises on computing *ker f* and working with the definition of *Normal*. However, the correctness of such materials is not established by the explanatory text itself; it is guaranteed only when all executable fragments are returned through *Π* to the kernel and re-checked by the predicate *K(Π(y))*. This is why even an AI-supported instructional module may remain mathematically rigorous.

7. Educational opportunities and limits of applicability

The proposed model has a direct didactic consequence: to every formally verified statement one may assign a learning object $\mathcal{U}(T, p)$, consisting of an explanation, a system of verified examples, a set of exercises, and hints. Only those components of such an object that are projected into executable kernel artefacts are mathematically significant; informal texts may improve intelligibility but do not participate in establishing the truth of the statement.

Therefore a natural requirement for a learning module is the inclusion $Exec(\mathcal{U}(T, p)) \subseteq \mathcal{P}ver$, where $\mathcal{P}ver$ is the set of all fragments of the kernel language that successfully pass type checking and thus belong to verified formal content. Under this condition, any example, proof template, or exercise with a code fragment preserves the same mathematical correctness as the original proof object *p*.

This also yields a methodological restriction: the generative component cannot be an independent arbiter of correctness; it can only serve as a means of explanation, reformulation, and structuring of already verified content. Accordingly, in the educational process it is expedient to distinguish three levels: verified kernel knowledge, the presentation interface, and generative support services.

It is precisely such a separation that makes it possible to construct courses for researchers, teachers, and students in which formal rigor is not lost in the transition to ordinary mathematical language. At the level of instructional design, this means the possibility of building verifiable examples, parameterized exercises, and systems of contextual hints for which the only criterion of correctness remains repeated verification of executable fragments by the Lean kernel.

8. Conclusions

The paper has developed a mathematical model of automatic verification of formalized proofs in Lean and a model of a conservative human-oriented presentation interface over its trusted kernel. The formalizable fragment of mathematical discourse, the kernel-verification predicate, a course formalizability functional, the elaboration and rendering model, and the safety condition for a generative extension have been defined.

It has been proved that, given adequate encoding, any statement from the set $\mathscr{ML}$ for which a proof object p has been constructed can be verified automatically. It has also been shown that ordinary mathematical notation, abbreviations, macros, and natural-language explanations do not enlarge the set of derivable statements whenever they are eliminated into kernel terms without changing the rules *K*.

A separate result is the formal safety condition for the generative component. It shows that generative technologies may be used to construct learning modules, lecture notes, explanations, examples, and assessment materials, but should not be regarded as a source of mathematical truth. The practical significance of the results lies in the possibility of building intelligent educational systems for courses in higher mathematics, discrete mathematics, algebra, and mathematical logic, in which the rigor of formalization is combined with an interface suitable for educational use.

REFERENCES

1. L. de Moura, S. Kong, J. Avigad, F. van Doorn, and J. von Raumer, “The Lean Theorem Prover (System Description),” in *Automated Deduction – CADE-25*, vol. 9195, 2015, pp. 378–388, https://doi.org/doi: 10.1007/978-3-319-21401-6_26.
2. L. de Moura and S. Ullrich, “The Lean 4 Theorem Prover and Programming Language,” in *Automated Deduction – CADE 28*, vol. 12699, 2021, pp. 625–635, doi: 10.1007/978-3-030-79876-5_37.
3. The mathlib Community, “The Lean Mathematical Library,” in *Proc. CPP 2020*, 2020, pp. 367–381, https://doi.org/doi: 10.1145/3372885.3373824.
4. T. Coquand and G. Huet, “The calculus of constructions,” *Inf. Comput.*, vol. 76, no. 2–3, pp. 95–120, 1988, doi: 10.1016/0890-5401(88)90005-3.
5. P. Martin-Löf, *Intuitionistic Type Theory*. Napoli, Italy: Bibliopolis, 1984.
6. P. Wadler, “Propositions as Types,” *Commun. ACM*, vol. 58, no. 12, pp. 75–84, 2015, https://doi.org/10.1145/2699407.
7. J. Avigad and P. Massot, *Mathematics in Lean*. Lean community tutorial, 2025. [Online]. Available: https://leanprover-community.github.io/mathematics_in_lean/
8. A. Thoma and P. Iannone, “Learning about Proof with the Theorem Prover LEAN: the Abundant Numbers Task,” *Int. J. Res. Undergrad. Math. Educ.*, vol. 8, pp. 64–93, 2022, doi: 10.1007/s40753-021-00140-1.
9. P. Iannone and A. Thoma, “Interactive theorem provers for university mathematics: an exploratory study of students’ perceptions,” *Int. J. Math. Educ. Sci. Technol*., 2024, https://doi.org/doi: 10.1080/0020739X.2023.2178981.
10. X. K. Yan and G. Hanna, “Using the Lean interactive theorem prover in undergraduate mathematics,” *Int. J. Math. Educ. Sci. Technol*., 2023, https://doi.org/doi: 10.1080/0020739X.2023.2227191.
11. P. Massot, “Teaching Mathematics Using Lean and Controlled Natural Language,” in *Leibniz Int. Proc. Inform. (LIPIcs)*, vol. 309, 2024, Art. no. 27, https://doi.org/doi: 10.4230/LIPIcs.ITP.2024.27.
12. J. Wemmenhove et al., “Waterproof: Educational Software for Learning How to Write Mathematical Proofs,” arXiv:2211.13513, 2022, https://doi.org/doi: 10.48550/arXiv.2211.13513.
13. O. Yevdokymov, A. Chukhray, and T. Stoliarenko, “Operationalizing the Formalizability of Mathematics Problems for Intelligent Tutoring Systems: Taxonomy, Measurement Protocol, and Educational Impact,” *CEUR Workshop Proc*., vol. 4164, paper 25, 2026.
14. A. Chukhray, D. Dvinskykh, V. Narozhnyy, and T. Stoliarenko, “Using an expression tree for adaptive learning,” in 2023 13th Int. Conf. Dependable Syst., Services Technol. (DESSERT), Athens, Greece, 2023, pp. 1–5, https://doi.org/10.1109/DESSERT61349.2023.10416497.
15. A. Chukhray, T. Stoliarenko, O. Yevdokymov, and V. Demyanenko, “Possibilities of using Intelligent Tutoring Systems (ITS) in Higher Mathematics Courses,” *Open Inf. Comput. Integr. Technol.*, no. 102, pp. 92–119, 2025, https://doi.org/10.32620/oikit.2024.102.07.

Євдокимов Олександр *аспірант, кафедра математичного моделювання та штучного інтелекту, Національний аерокосмічний університет «Харківський авіаційний інститут», вул. Вадима Манька, 17, Харків, Україна, 61070*
e-mail: o.yevdokymov@khai.edu
https://orcid.org/0009-0008-9687-6344

Лучшева Оксана *старший викладач, кафедра інженерії програмного забезпечення, Національний аерокосмічний університет «Харківський авіаційний інститут», вул. Вадима Манька, 17, Харків, Україна, 61070*
e-mail: o.luchsheva@khai.edu ;
https://orcid.org/0000-0003-3855-2815

Математична модель автоматичної верифікації формалізованих доказів і консервативний інтерфейс представлення в Lean

Актуальність. Інтерактивні системи доведення теорем, такі як Lean, дозволяють перенести остаточний контроль математичної правильності від людини до малого довіреного ядра. Однак їх широке застосування в дослідженнях і вищій освіті все ще обмежене значним розривом між строгим формальним мовою ядра, заснованою на залежних типах, та звичною інтуїтивною математичною нотацією, символами й міркуваннями природною мовою, якими користуються математики та студенти. Цей розрив створює серйозні перешкоди для навчання доведенню, формалізації нових результатів і розробки ефективних інтелектуальних освітніх систем.
Мета. Метою роботи є побудова строгої математичної моделі автоматичної верифікації формалізованих доказів у Lean та надання формального обґрунтування можливості створення консервативного інтерфейсу, орієнтованого на користувача, разом з ненадійним генеративним педагогічним компонентом над довіреним ядром без втрати математичної строгості.
Методи. Дослідження ґрунтується на теорії залежних типів, метатеоретичних властивостях обчислення конструкцій (коректність, сильна нормалізація, розв'язність перевірки типів), моделях розробки та зворотної проекції між мовою представлення і мовою ядра, а також формальних предикатах правильності типізації термінів. Безпека генеративного AI-компонента визначається умовою, що кожен виконуваний артефакт має бути спроектований назад у ядро і повторно перевірений довіреним механізмом перевірки типів.
Результати. У розробленій моделі визначено формалізований фрагмент математичного дискурсу $\mathscr{M}$L, предикат верифікації ядра VerL, функціонал формалізованості курсу $\mathbb{F}(\mathcal{S})$, повну модель консервативного інтерфейсу та умову безпеки генеративного розширення. Доведено, що синтаксичний цукор, макроси, скорочення нотацій та AI-генеровані освітні артефакти не розширюють множину виводних тверджень, якщо довірене ядро Lean залишається незмінним і всі виконувані фрагменти повторно перевіряються ядром. Ілюстративний приклад з чотирма теоремами з алгебри та теорії груп демонструє, що нотація, орієнтована на користувача, повністю розгортається в терміни ядра зі збереженням математичного змісту та перевірюваності. Показник вірності інтерфейсу на тестовому наборі досягає максимального значення 1.
Висновки. Отримані результати демонструють, що формальну мову Lean можна оснастити зручним інтерфейсом, орієнтованим на користувача, та інтелектуальними освітніми сервісами, придатними для наукового та навчального використання, без втрати математичної строгості. Єдиним критерієм істинності залишається повторна верифікація малим довіреним ядром. Запропоноване розділення між верифікованими знаннями ядра, консервативним шаром представлення та ненадійними генеративними сервісами створює надійну основу для побудови інтелектуальних освітніх систем з вищої математики, дискретної математики, алгебри та математичної логіки, у яких формальна правильність гарантується незалежно від якості згенерованих пояснень і вправ.

Ключові слова: *Lean, залежні типи, автоматична верифікація доказів, формалізована математика, консервативний інтерфейс, генеративний педагогічний модуль.*

Надійшла у першій редакції 20.01.2026; рекомендована до друку 04.03.2026; опублікована 03.04.2026.
The first version has been received on 20.01.2026; was recommended for printing on 04.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-04

УДК (UDC) 004.93

Коршенко Владислав Сергійович *Аспірант кафедри Кафедри кібербезпеки інформаційних систем, мереж і технологій, старший викладач кафедри математичного моделювання та аналізу даних, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022*
e-mail: v.korshenko@karazin.ua
https://orcid.org/0000-0003-2197-072X

Узлов Дмитро Юрійович *Кандидат технічних наук, Директор ННІ КН та ШІ, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022*
e-mail: dmytro.uzlov@karazin.ua*;*
https://orcid.org/0000-0003-3308-424X

Оцінка впливу наявності фотореалістичної текстури при генерації синтетичного датасету на точність моделей комп'ютерного зору

Актуальність. Сучасний розвиток комп'ютерного зору стикається з проблемою високої вартості та трудомісткості збору реальних анотованих даних. Використання синтетичних даних, згенерованих у графічних рушіях, є ефективною альтернативою, проте головною перешкодою залишається «розрив між доменами» (domain gap), що знижує точність моделей на реальних зображеннях.
Метою роботи є кількісна оцінка впливу фотореалістичної текстури цільового об'єкта на ефективність детектування моделями YOLO при переході від симуляції до реальності (Sim2Real).
Методологія дослідження базується на проведенні контрольованого експерименту в середовищі Unity, де було згенеровано два ідентичні синтетичні датасети, що відрізнялися лише типом текстури 3D-моделі: високодеталізованою фотореалістичною («Textured») та монохромною білою («White»). Навчання моделей проводилося на базі архітектури YOLOv11s із застосуванням стратегії переносу навчання (transfer learning) та двоетапного процесу тонкого налаштування. Валідація результатів здійснювалася на незалежному наборі виключно реальних фотографій.
Результати. Обидві моделі, що були навчені на двох датасетах («Textured» і «White»), досягли майже ідентичної точності на синтетичних валідаційних даних (mAP@0.5 ≈ 0.995). Однак на реальних фотографіях модель «Textured» продемонструвала в 11.6 разів вищий mAP@0.5, порівняно з результатом моделі «White». Показник повноти (recall) для текстурованої моделі виявився в 10.3 рази вищим, ніж у моделі, що покладалася лише на геометричну форму.
Висновки. Фотореалістична текстура є критично важливим чинником для успішного Sim2Real перенесення. Вона забезпечує формування в ранніх шарах нейронної мережі універсальних низькорівневих ознак, які є необхідними для розпізнавання об'єктів у реальному середовищі. Якісне текстурування 3D-асетів слід розглядати як стратегічний пріоритет, а не допоміжний етап візуалізації.

Ключові слова: *синтетичні дані, комп'ютерний зір, детектування об'єктів, розрив між доменами, робастність моделей, стійкість до зсуву домену.*

Як цитувати: Коршенко В. С., Узлов Д. Ю., "Оцінка впливу наявності фотореалістичної текстури при генерації синтетичного датасету на точність моделей комп'ютерного зору". *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.41-58. https://doi.org/10.26565/2304-6201-2026-69-04

How to quote: V Korshenko, D.Uzlov, "Assessment of the impact of photorealistic textures on the accuracy of computer vision models using synthetic datasets", *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical Modelling. Information Technology. Automated Control Systems*, vol. 69, pp. 41–58, 2026. https://doi.org/10.26565/2304-6201-2026-69-04 [in Ukrainian]

Вступ

Сучасні досягнення в галузі комп'ютерного зору, особливо в задачах детектування об'єктів та семантичної сегментації, нерозривно пов'язані з використанням глибоких нейронних мереж, які для свого навчання потребують великих обсягів якісно анотованих даних [1].

Проте процес збору та ручної розмітки реальних зображень є надзвичайно трудомістким, фінансово витратним і часто пов'язаний з логістичними або етичними обмеженнями [2].

Ці виклики стають особливо гострими при роботі з рідкісними сценаріями – наприклад, аварійними ситуаціями для безпілотних транспортних засобів, або в умовах, де збір даних є небезпечним чи неможливим [11].

У відповідь на ці проблеми, все більшої популярності набуває використання **синтетичних даних**, згенерованих за допомогою комп'ютерної графіки [1].

Сучасні графічні рушії, такі як *Unity* та *Unreal Engine*, дозволяють створювати фотореалістичні сцени, які містять повністю контрольовані параметри освітлення, текстур і матеріалів, а також автоматично генерують розмітку (обмежувальні прямокутники, маски сегментації, карти глибини) [10].

Цей підхід знімає обмеження, пов'язані з браком даних, забезпечує стабільність якості анотацій та дозволяє масштабувати експерименти без втрати контрольованості [3].

Зростання цього напряму підтверджується академічними аналізами, які вказують на стрімке розширення ринку та застосувань синтетичних даних [18].

Однак, незважаючи на переваги, ключовою проблемою залишається так званий **розрив між доменами** (*domain gap* або *Sim2Real gap*), який проявляється у зниженні точності моделей при переході з симуляційних до реальних даних [12].

Цей розрив зумовлений статистичними та візуальними відмінностями між синтетичними та реальними зображеннями – від текстур і шуму сенсорів до складності освітлення та фону [13].

Для подолання цієї проблеми наукова спільнота розробила три основні стратегії [6]:

1. Доменна рандомізація (Domain Randomization, DR) – варіювання текстур, освітлення, ракурсів і матеріалів у широкому (навіть не реалістичному) діапазоні, що дозволяє моделі навчитися узагальнювати незалежно від конкретних умов [3], [4];
2. Доменна адаптація (Domain Adaptation, DA) – статистичне узгодження розподілів ознак між доменами з використанням нейронних мереж або змагальних методів [5], [6];
3. Підвищення фотореалізму (Photorealism) – фізично коректне відтворення текстур, матеріалів і освітлення для зменшення візуальної різниці між симуляційним і реальним світом [7], [8].

Саме третій підхід є фокусом цього дослідження.

Попередні роботи Hinterstoisser et al. [7] продемонстрували, що використання високоякісних, фотореалістичних синтетичних даних може забезпечити продуктивність, наближену до результатів навчання на реальних вибірках. Водночас внесок окремих аспектів реалізму – зокрема наявності або відсутності текстури об'єкта – залишається недостатньо вивченим.

Дослідження Jackson et al. [4] показують, що навіть у межах доменної рандомізації складність текстур підвищує точність моделей, що вказує на ключову роль текстурної інформації у формуванні переносимих ознак.

Теоретичну основу цього припущення заклали Yosinski et al. [8], які експериментально довели, що ранні шари згорткових нейронних мереж навчаються розпізнавати універсальні низькорівневі ознаки (градієнти, краї, кольорові переходи), тоді як пізніші – спеціалізовані, залежні від конкретної задачі.

Це означає, що саме якісна текстуризація синтетичних об'єктів може покращувати переносимість моделі при переході між доменами, оскільки збагачує її вхідні сигнали низькорівневими характеристиками, спільними для обох середовищ.

Метою даної роботи є експериментальна та кількісна оцінка впливу фотореалістичної текстури цільового об'єкта в синтетичних датасетах на ефективність детектування об'єктів моделлю архітектури YOLO[10] при переході від синтетичних до реальних даних (Sim2Real).

Для досягнення поставленої мети було розроблено контрольований експеримент, у якому варіювався виключно фактор текстуризації 3D-моделі, тоді як усі інші параметри сцени, процесу навчання та валідації залишалися незмінними.

Висунута гіпотеза полягає в тому, що саме фотореалістичні текстури є одним із ключових чинників зменшення розриву *Sim2Real*, підвищуючи точність та узагальнюваність моделей комп'ютерного зору.

2. Матеріали та методи

Цей розділ детально описує методологію, використану для перевірки нашої дослідницької гіпотези. Експериментальний дизайн було розроблено таким чином, щоб забезпечити контрольовані умови для ізоляції та кількісної оцінки впливу фотореалістичної текстури на ефективність моделі комп'ютерного зору. Процес дослідження було послідовно розділено на три основні етапи: (1) генерація двох варіантів синтетичного датасету, що відрізняються лише однією цільовою характеристикою; (2) навчання двох ідентичних моделей детектування об'єктів на цих датасетах за однакових умов; (3) валідація та порівняльний аналіз продуктивності моделей на незалежних тестових наборах, що містять реальні зображення.

Основою нашого експерименту було створення висококонтрольованих синтетичних наборів даних. Цей підхід дозволив нам ізолювати вплив текстури об'єкта як єдиної змінної, усунувши сторонні фактори, які могли б вплинути на результати навчання моделі. Для генерації зображень було обрано ігровий рушій **Unity**, оскільки він надає широкий набір інструментів для створення фотореалістичних сцен, симуляції освітлення та автоматизації процесу збору даних з ідеально точною розміткою.

Для створення віртуального середовища ми використали готову 3D-сцену лісу (Рис. 1), використану на підставі ліцензії Unity Asset Store та її умов використання. Ця сцена створювала оптимальний візуальний контекст з природним оточенням, що містить різноманітні фонові об'єкти (дерева, кущі, траву) та неоднорідний ландшафт, що дозволило згенерувати складні для аналізу зображення.

Рис. 1: Сцена лісу, що була використана як оточення для розміщення моделі цільового об’єкту
Figure 1: Forest scene used as the environment for placing the target object model

В якості цільового об'єкта для детектування було обрано 3D-модель качки (Рис. 2). Цей вибір був зумовлений трьома основними причинами. По-перше, з точки зору геометрії, об'єкт має помірну складність: він поєднує плавні контури та вигнуті поверхні, але не має надмірно дрібних деталей, таких як хутро чи окремі пера, що могло б ускладнити аналіз впливу саме базової текстури. По-друге, різноманітність природних забарвлень качок робить їх чудовим прикладом для вивчення важливості текстурних ознак. По-третє, наявність високодеталізованих 3D-моделей у вільному доступі спростила підготовку до експерименту. Модель було розміщено у центрі сцени, щоб забезпечити достатній простір для позиціонування камери з різних ракурсів.

Рис. 2: Сцена з розміщеним у центрі цільовим об’єктом
Figure 2: Scene with a target object placed in the center

Освітлення в сцені було реалізовано за допомогою одного глобального джерела світла типу Directional Light, що імітує сонячне світло. Для досягнення м'яких та реалістичних тіней було встановлено параметр Shadow Type у значення Soft Shadows, а інтенсивність світла Intensity – 2.27. Налаштування оточення, такі як матеріал неба (Skybox) та загальне розсіяне світло (Ambient Color), також були сконфігуровані для створення природної денної атмосфери. Від цих параметрів залежали ключові візуальні аспекти сцени: яскравість, контрастність, колір тіней та відблиски на поверхнях.

Збір зображень виконувався за допомогою віртуальної камери (Рис. 3), позиція якої контролювалася спеціально розробленим скриптом. Цей скрипт дозволив автоматизувати процес зйомки, систематично переміщуючи камеру навколо цільового об'єкта. Камера оберталася на фіксованій відстані (Distance = 35 одиниць) з визначеним кроком по горизонталі (Horizontal Step = 20 градусів) та вертикалі (Vertical Step = 30 градусів). Такий підхід гарантував, що об'єкт буде знято з великої кількості різноманітних, але відтворюваних ракурсів.

Рис. 3: Віртуальна камера, розміщена на сцені, та її поле зору
Figure 3: Virtual camera placed on the scene and its field of view

Головною умовою експерименту було створення двох наборів даних, які б відрізнялися виключно текстурою цільового об'єкта. Усі інші параметри – геометрія 3D-моделі, налаштування сцени, освітлення, траєкторія руху камери – залишалися абсолютно ідентичними для обох випадків.

Датасет 1 - "Textured" (Верхня половина Рис. 4): Для цього набору даних на 3D-модель качки було накладено високодеталізовану, фотореалістичну текстуру, яка імітувала природне забарвлення птаха.

Датасет 2 - "White" (Нижня половина Рис. 4): У цьому випадку та сама 3D-модель використовувалася з простою монохромною білою текстурою. Вона не несла жодної інформації про колір чи патерни, дозволяючи моделі спиратися лише на геометричну форму об'єкта, його силует та затінення.

Такий підхід дозволив нам створити дві паралельні реальності для навчання, де єдиною відмінністю була наявність або відсутність текстурних ознак.

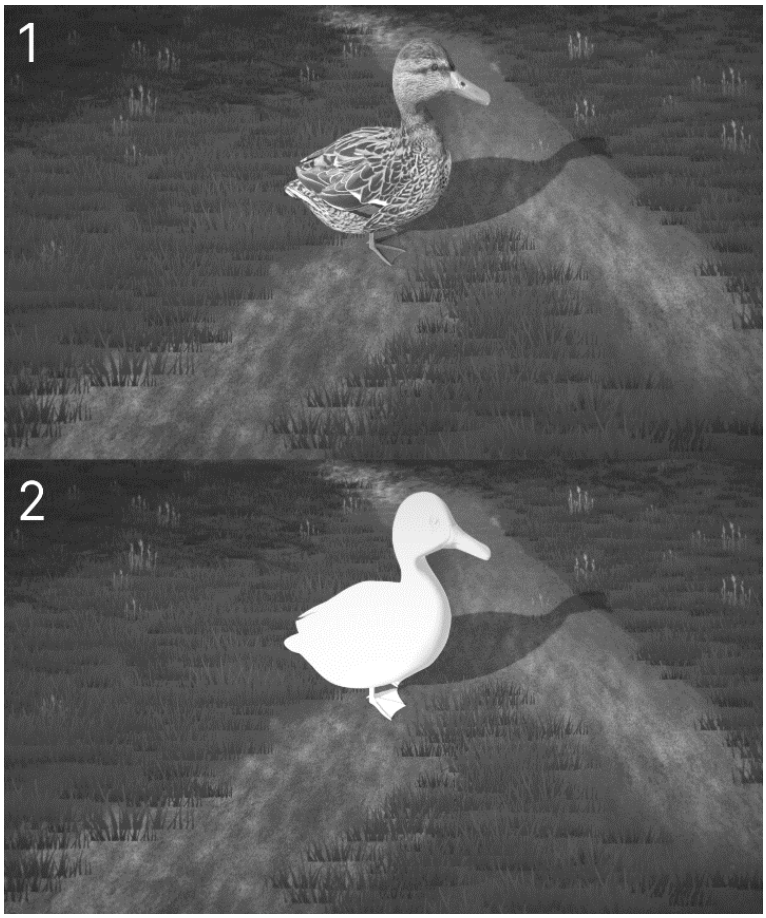


Рис. 4: Побічне порівняння ідентичних кадрів з датасету "Textured" (вгорі) та "White" (внизу)
Figure 4: Side-by-side comparison of identical frames from the "Textured" (top) and "White" (bottom) datasets

Процес розмітки даних (створення обмежувальних прямокутників, або bounding boxes) був повністю автоматизований засобами Unity. Для кожного згенерованого кадру скрипт розраховував точні екранні координати видимої частини цільового об'єкта. Це досягалося шляхом випускання променів (RayCast) з камери до кожного пікселя екрана; якщо промінь перетинав 3D-модель об'єкта, його координати використовувалися для визначення крайніх точок (верхньої, нижньої, лівої, правої), на основі яких і будувався обмежувальний прямокутник . Координати зберігалися у текстовому файлі у форматі, сумісному з YOLO.

Для кожного з двох сценаріїв ("Textured" та "White") було згенеровано по 342 унікальних зображення. Отримані дані були розділені на навчальну та валідаційну вибірки у співвідношенні приблизно 80/20:

- 273 зображення для навчання.
- 69 зображень для валідації.

Розподіл було здійснено таким чином, щоб зображення з однакових ракурсів в обох датасетах потрапляли у відповідні вибірки (навчальну до навчальної, валідаційну до валідаційної), що забезпечило повну узгодженість умов для подальшого навчання та порівняння моделей.

2.1 Архітектура та навчання моделі детектування об'єктів

Для об'єктивного порівняння впливу текстури було критично важливо, щоб обидві моделі ("textured" та "white") були ідентичними за архітектурою та навчалися за абсолютно однакових умов. Цей підрозділ описує вибір моделі, стратегію її навчання та конкретні гіперпараметри, що використовувалися для забезпечення відтворюваності експерименту.

В якості базової архітектури для вирішення задачі детектування об'єктів було обрано модель YOLOv11s з репозиторію Ultralytics. Ця модель є представником сімейства YOLO (You Only Look Once), яке добре відоме своїм балансом між швидкістю роботи та точністю. Конфігурація "s" (small) є полегшеною версією, що містить близько 7 мільйонів параметрів. Такий вибір був зумовлений прагматичними міркуваннями: малий обсяг ваг моделі значно знижує вимоги до обчислювальних ресурсів, зокрема до обсягу відеопам'яті (VRAM), що дозволило проводити серію експериментів на GPU середнього класу та прискорити ітерації дослідження.

Навчання глибокої нейронної мережі «з нуля» на нашому відносно невеликому навчальному наборі з 273 синтетичних зображень неминуче призвело б до сильного перенавчання (*overfitting*), коли модель ідеально запам’ятовує навчальні приклади, але втрачає здатність до узагальнення на нових даних. Щоб уникнути цієї проблеми, ми застосували стратегію переносу навчання (*transfer learning*).

Ми використали модель YOLOv11s, попередньо навчену на великому та різноманітному датасеті COCO (Common Objects in Context), який містить мільйони об’єктів різних класів.

Згідно з Yosinski et al. [8], ранні шари згорткових мереж формують універсальні низькорівневі ознаки (градієнти, контури, текстури), які легко переносяться між доменами.

Це пояснює, чому використання попередньо натренованих ваг YOLOv11, отриманих на великій базі COCO, сприяє кращому *transfer learning* під час навчання на синтетичних даних.

Така здатність до адаптації є ключовою для зменшення розриву *Sim2Real* та забезпечує узгодженість результатів між симульованим і реальним середовищами.

Цей підхід ґрунтується на фундаментальному спостереженні, що ранні згорткові шари нейронної мережі навчаються розпізнавати загальні низькорівневі та середньорівневі ознаки, такі як краї, градієнти, кольорові плями та базові текстури. Ці ознаки є універсальними і корисними для широкого спектра задач комп’ютерного зору. Натомість, більш глибокі шари, що знаходяться ближче до виходу мережі, спеціалізуються на виявленні високоспецифічних ознак, що стосуються конкретних класів із початкового датасету (наприклад, COCO).

Таким чином, стратегія полягала в тому, щоб зберегти корисні універсальні фільтри з «кістяка» (*backbone*) моделі та адаптувати лише її «голову» (*head*) до нашої специфічної задачі – детектування одного класу «качка».

Для забезпечення стабільного та ефективного навчання було розроблено двоетапну стратегію, яка поєднувала "заморозку" шарів та тонке налаштування.

Етап 1: "Розігрів" голови детектора (Head Warm-up). На цьому етапі метою було адаптувати лише вихідні, класифікаційні шари моделі до нового завдання, не зачіпаючи стабільні ваги попередньо навченого кістяка. Для цього перші 10 шарів нейронної мережі були "заморожені" (параметр freeze=10), тобто їхні ваги не оновлювалися під час градієнтного спуску. Навчання тривало протягом 5 епох з відносно високою швидкістю навчання (learning rate) 10^{-4}. Такий підхід, рекомендований Ultralytics, дозволяє швидко адаптувати модель до нового класу без ризику руйнування універсальних фільтрів.

Етап 2: Тонке налаштування всієї мережі (Full Network Fine-tuning). Після того, як "голова" моделі була адаптована, ми "розморозили" всю мережу і продовжили її донавчання протягом 80 епох. На цьому етапі швидкість навчання було значно знижено до $3*10^{-5}$. Мета цього етапу – дозволити всім шарам моделі, включно з ранніми, тонко підлаштуватися під специфіку наших синтетичних даних, зберігаючи при цьому стабільність градієнтів завдяки низькій швидкості навчання. Вибір такої комбінації (короткий "розігрів" + тривале тонке налаштування) спирався на результати досліджень, які показали, що такий підхід підвищує стабільність та кінцеву точність моделі, особливо в умовах значного розриву між доменами.

В Таблиці 1 наведено параметри доетапного навчання.

Таблиця 1. Параметри двоетапного навчання
Table 1. Parameters of two-stage training

Етап	Шари для навчання	Епохи	Швидкість навчання (LR)	Заморожені шари (Freeze)	Мета
1	Головні детекторні head-шари	5	$1*10^{-4}$	0-9	Адаптувати класифікатор до класу duck без руйнування універсальних фільтрів.
2	Уся мережа	80	$3*10^{-5}$	–1	Тонко підлаштувати усі рівні, зберігаючи стабільність градієнтів.

Для гарантування повної відтворюваності результатів усі експерименти проводилися у детермінованому режимі.

Було зафіксовано початкове значення генератора псевдовипадкових чисел (seed = *42*) для всіх етапів навчання, що забезпечило ідентичну ініціалізацію ваг і послідовність подачі даних при кожному запуску моделі.

Такий підхід відповідає базовим принципам наукової відтворюваності в машинному навчанні, описаним у роботі Picard [19], де наголошується на важливості контролю стохастичних процесів для коректного порівняння результатів між експериментами.

У ході навчання застосовувались такі основні гіперпараметри:

- Batch size: 32
- Learning rate: $1*10^{-4}$ (етап 1) / $3*10^{-5}$ (етап 2)
- Оптимізатор: Adam [20]
- Функція втрат: комбінація *Binary Cross-Entropy* та *Complete IoU Loss (CIoU)* [21]
- Розмір вхідних зображень: 640×640 пікселів
- Freeze: 10 шарів на етапі попередньої адаптації
- Epochs: 5 + 80 (двоступеневе навчання)

Оптимізатор Adam (Adaptive Moment Estimation) було обрано через його ефективність і швидку збіжність при донавчанні моделей на невеликих наборах даних.

Як показано у роботі Kingma та Ba [20], цей алгоритм адаптивно коригує швидкість навчання для кожного параметра, поєднуючи переваги AdaGrad і RMSProp, що забезпечує швидшу стабілізацію градієнтів.

Функція втрат поєднує *Binary Cross-Entropy (BCE)* для класифікаційної складової та *Complete IoU Loss (CIoU)* для регресії обмежувальних рамок.

Згідно з дослідженням Zheng et al. [21], CIoU враховує не лише площу перетину рамок, а й відстань між їх центрами та співвідношення сторін, що дозволяє моделі ефективніше локалізувати об’єкти в кадрі порівняно зі звичайним IoU.

Додатково використовувалась стратегія плавного зниження швидкості навчання за косинусним законом (Cosine Annealing Scheduler) [22], реалізована параметром cos_lr=True.

Метод був уперше описаний у роботі Loshchilov і Hutter [22] і довів свою ефективність у запобіганні «перестрибуванню» через локальні мінімуми, забезпечуючи більш плавну збіжність і кращу генералізацію на пізніх етапах навчання.

Також застосовувався механізм «прогріву» (Warm-up) [23], протягом перших двох епох (warmup_epochs=2), коли швидкість навчання поступово збільшувалася від нуля до встановленого значення.

Як показано у роботі Goyal et al. [23], такий підхід стабілізує оптимізацію, особливо під час початкових етапів тренування моделей із великим розміром батчу.

У сукупності ці процедури – детермінізація, адаптивна оптимізація, комбінована функція втрат, плавна зміна швидкості навчання та етап прогріву – забезпечили стабільну збіжність, стійкість до коливань градієнтів і відтворюваність результатів експериментів.

2.2 Експериментальна валідація та метрики оцінювання

Після завершення навчання двох моделей – "textured" та "white" – наступним кроком стала їхня всебічна валідація та порівняльний аналіз. Для отримання об'єктивних та надійних результатів було розроблено методологію оцінювання, яка включала використання двох спеціально підготовлених тестових наборів даних та набір стандартних метрик якості для задач детектування об'єктів.

Щоб оцінити продуктивність моделей у різних умовах, ми підготували два незалежних тестових датасети, кожен з яких мав свою специфічну мету.

Тестовий датасет 1 - змішаний набір для базової оцінки. Цей набір містив 30 синтетичних зображень, взятих порівну з валідаційних вибірок обох навчальних датасетів, та 12 реальних фотографій цільових об'єктів (качок).

Основна мета цього датасету – виконати "перевірка адекватності" (sanity check). По-перше, оцінка на знайомих синтетичних даних дозволила переконатися, що обидві моделі успішно засвоїли свої відповідні домени ("in-domain performance"). По-друге, невелика кількість реальних зображень дала змогу провести первинну, "м'яку" перевірку здатності моделей до узагальнення (generalization) та перенесення знань на реальний світ.

Тестовий датасет 2 - реальний набір для повноцінної Cross-Domain валідації. Цей набір даних був розроблений для жорсткого тестування моделей в умовах, максимально наближених до практичного застосування, і повністю складався з реальних фотографій.

Він містив:

- 50 позитивних прикладів - реальні фотографії качок у різних середовищах, позах та умовах освітлення.
- 50 негативних прикладів, які, у свою чергу, були поділені на дві групи для перевірки стійкості до різних типів помилок:
- 25 "важких" негативних прикладів: фотографії інших птахів (гусей, лебедів, чайок), які мають схожий силует або знаходяться у схожому контексті. Це дозволило оцінити здатність моделей відрізняти цільовий клас від схожих нецільових об'єктів.
- 25 "легких" негативних прикладів: фотографії природних пейзажів (ліс, озеро) без будь-яких птахів чи тварин. Ця група була призначена для перевірки моделей на схильність до хибних спрацювань (false positives) на складних фонах.

Цей датасет є основним інструментом для відповіді на наше дослідницьке питання. Оцінка на ньому дозволяє виміряти реальну ефективність перенесення знань з симуляції в реальність (Sim2Real), а також оцінити надійність та практичну придатність кожної з моделей.

Для кількісної оцінки та порівняння продуктивності навчених моделей було використано набір стандартних метрик для задач детектування об'єктів. Якість детекції оцінювалася за метриками Precision (Точність), Recall (Повнота) та інтегрального показника mean Average Precision (mAP).

Для всебічного аналізу ми розраховували дві варіації mAP, що відрізняються за вимогами до точності локалізації:

- mAP@0.5 (mAP50): Ця метрика оцінює здатність моделі правильно класифікувати та в цілому локалізувати об'єкт, використовуючи поріг Intersection over Union (IoU) 0.5.
- mAP@0.5-0.95 (стандартна метрика COCO): Ця більш сувора метрика оцінює точність обмежувальних прямокутників, усереднюючи mAP по діапазону порогів IoU від 0.5 до 0.95.

Такий набір метрик дозволив нам провести комплексний аналіз, оцінивши як загальну здатність моделей до виявлення об'єктів, так і точність їхньої локалізації.

3. Результати

У цьому розділі представлено кількісні результати експериментів, проведених згідно з методологією, описаною в попередньому розділі. Дані викладено об'єктивно та послідовно, щоб продемонструвати ефективність обох навчених моделей – "textured" та "white" – на різних етапах валідації. Спочатку наведено результати навчання на синтетичних даних, що встановлюють базовий рівень продуктивності моделей у межах їхніх навчальних доменів. Далі представлено порівняльні результати тестування на змішаному та повністю реальному датасетах, що дозволяє оцінити здатність кожної моделі до перенесення знань та її ефективність в умовах зсуву домену.

3.1. Результати навчання та валідації на синтетичних даних

Першочерговим етапом аналізу була перевірка якості навчених моделей на валідаційних вибірках, що походили з тих самих синтетичних доменів, що й навчальні дані. Цей крок був необхідний для того, щоб встановити базовий рівень продуктивності та переконатися, що обидві моделі – "textured" та "white" – успішно засвоїли візуальні закономірності своїх відповідних наборів даних.

Результати внутрішньодоменної валідації (in-domain validation) продемонстрували, що обидві моделі досягли майже ідеальних та практично ідентичних показників. Для обох конфігурацій інтегральний показник mAP@0.5 склав приблизно 0.995. Це свідчить про те, що ні недонавчання (underfitting), ні суттєві відмінності в ємності засвоєної інформації не були факторами в експерименті. Обидві моделі повністю опанували свої синтетичні домени, що створило надійну основу для подальшого порівняння їхньої здатності до перенесення знань на реальні дані.

Висока якість навчання підтверджується візуальними метриками. Нормалізовані матриці сплутування для обох моделей демонструють 100% точність на валідаційній вибірці, де всі об'єкти класу "качка" були правильно ідентифіковані, а хибні спрацювання на фоні були відсутні (Рис. 5, Рис. 6).

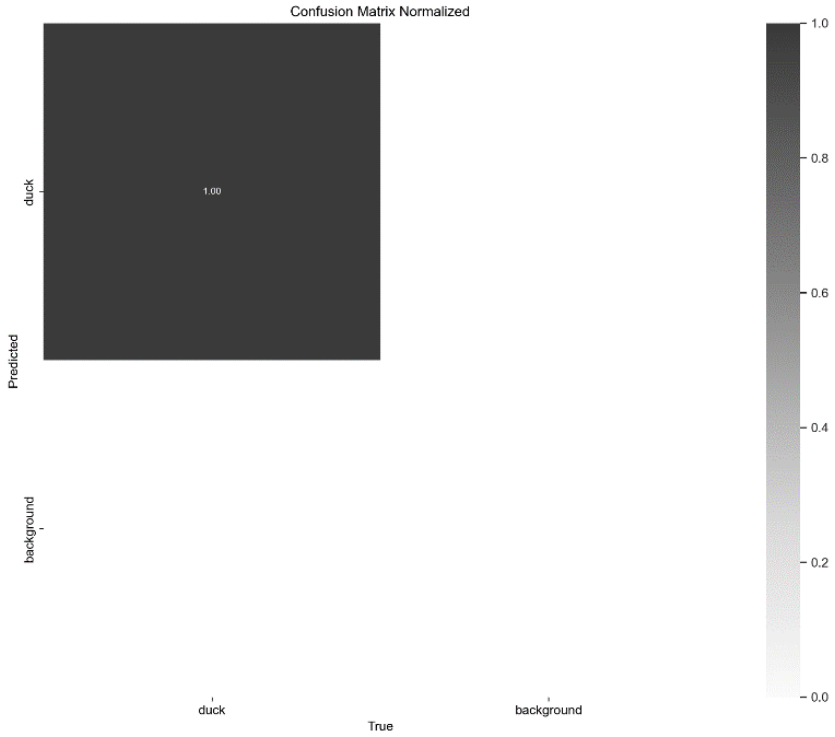


Рис. 5: Нормалізована матриця сплутування для моделі "textured" на синтетичному валідаційному датасеті

Figure 5: Normalized confusion matrix for the "textured" model on a synthetic validation dataset

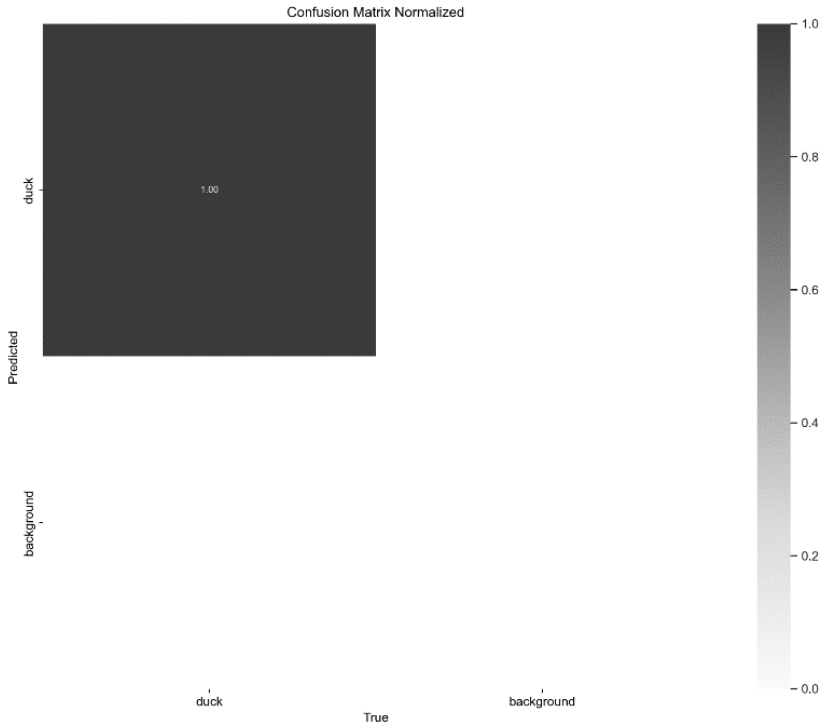


Рис. 6: Нормалізована матриця сплутування для моделі "white" на синтетичному валідаційному датасеті
Figure 6: Normalized confusion matrix for the "white" model on a synthetic validation dataset

Аналогічно, криві точності-повноти (Precision-Recall curves) для обох моделей мають прямокутну форму (Рис. 7, 8), що є характерним для детектора з дуже високою продуктивністю, де і точність, і повнота близькі до 1.0 у широкому діапазоні порогів впевненості.

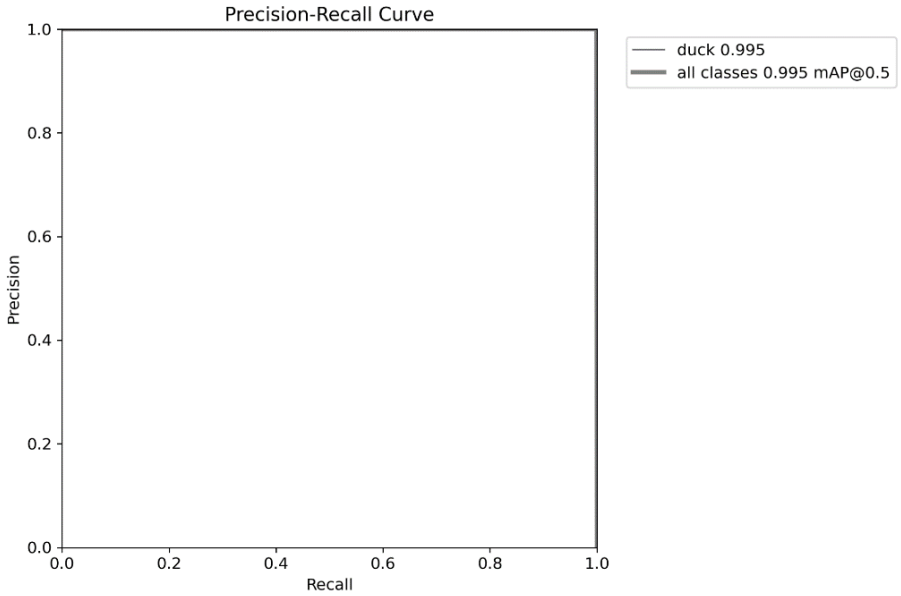


Рис. 7: Крива PR для моделі "textured" на синтетичному валідаційному датасеті
Figure 7: PR curve for the "textured" model on a synthetic validation dataset

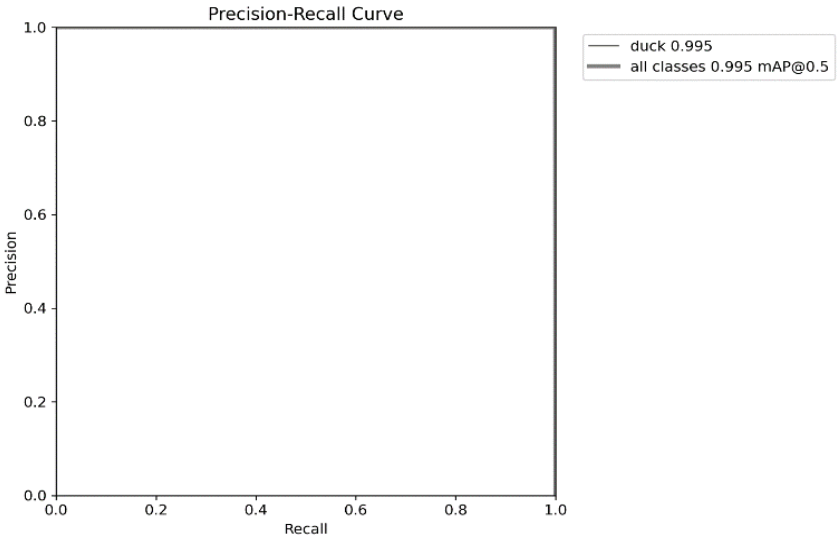


Рис. 8: Крива PR для моделі "white" на синтетичному валідаційному датасеті
Figure 8: PR curve for the "white" model on a synthetic validation dataset

Таким чином, результати цього етапу підтвердили, що обидві моделі були успішно та стабільно навчені за однакових умов, що дозволяє впевнено стверджувати, що будь-які подальші розходження в їхній продуктивності на реальних даних будуть зумовлені виключно впливом досліджуваного фактора – наявністю або відсутністю фотореалістичної текстури.

3.2. Порівняльна оцінка на змішаному тестовому датасеті

Наступним етапом було тестування обох моделей на змішаному датасеті, що складався як із синтетичних, так і з реальних зображень. Цей експеримент дозволив оцінити продуктивність моделей в умовах "м'якого" зсуву домену, коли детектор стикається з першими прикладами з реального світу.

Модель, навчена на даних з фотореалістичною текстурою, продемонструвала досить високу продуктивність. Показник mAP@0.5 досяг 0.682, що свідчить про спроможність моделі узагальнювати отримані знання. Показник точності (precision) залишився на високому рівні 0.965, однак повнота (recall) знизилася до 0.534. Це вказує на те, що модель, хоч і робила впевнені та правильні детекції, пропускала майже половину цільових об'єктів на реальних зображеннях.

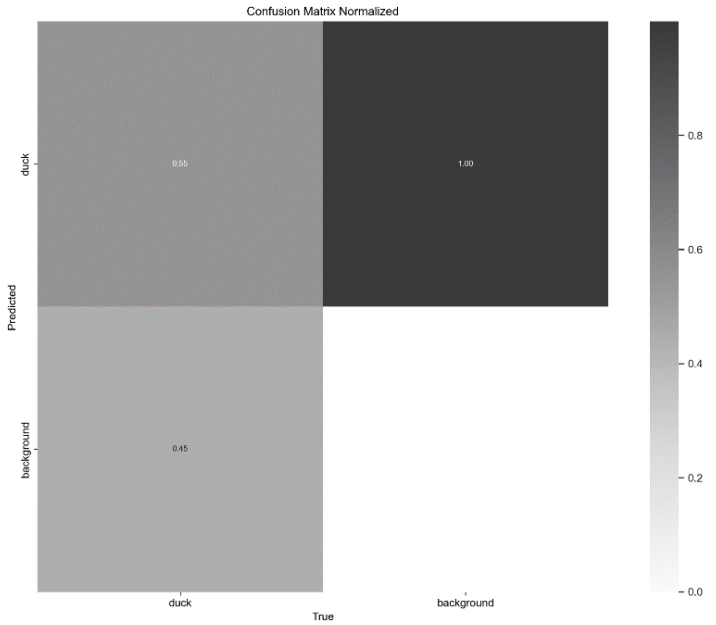


Рис. 9: Нормалізована матриця сплутування для моделі "textured" на змішаному датасеті
Figure 9: Normalized confusion matrix for the "textured" model on a mixed dataset

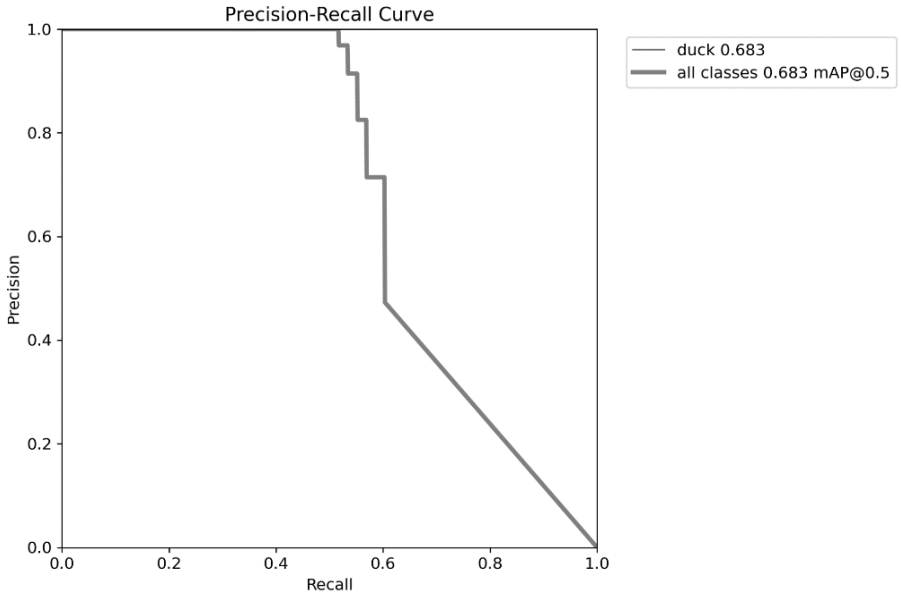


Рис. 10: Крива PR для моделі "textured" на змішаному датасеті
Figure 10: PR curve for the "textured" model on a mixed dataset

Модель, навчена на даних з монохромною білою текстурою, показала значно нижчі результати за всіх ключових метриках. Показник mAP@0.5 склав лише 0.421, а більш суворий mAP@0.5-0.95 – 0.359. Особливо помітним було падіння повноти (recall), яка становила всього 0.310. Це означає, що модель, яка спиралася переважно на геометричну форму, змогла правильно ідентифікувати менше третини цільових об'єктів у нових умовах.

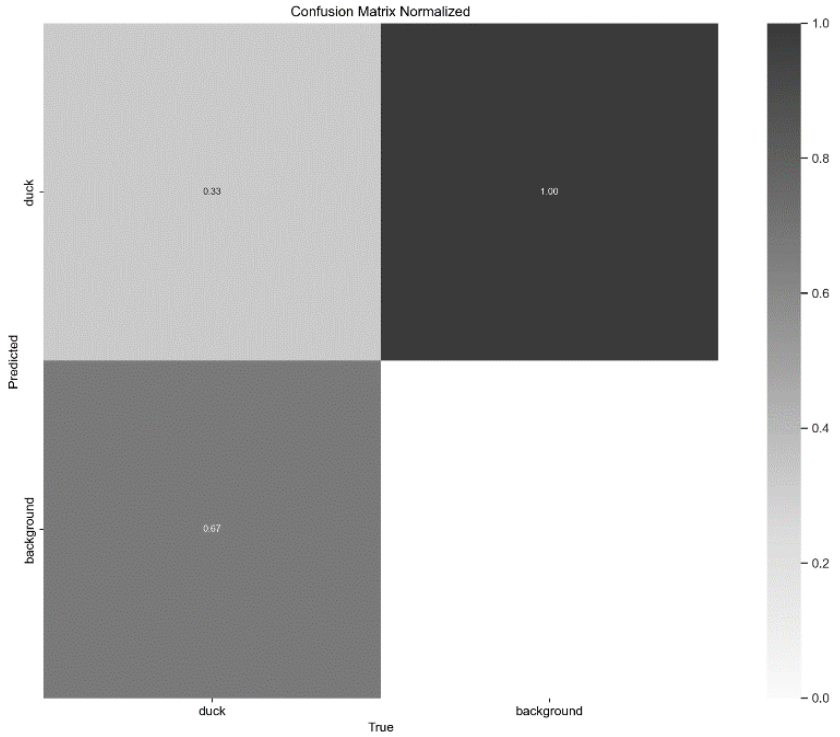


Рис. 11: Нормалізована матриця сплутування для моделі "white" на змішаному датасеті
Figure 11: Normalized confusion matrix for the "white" model on a mixed dataset

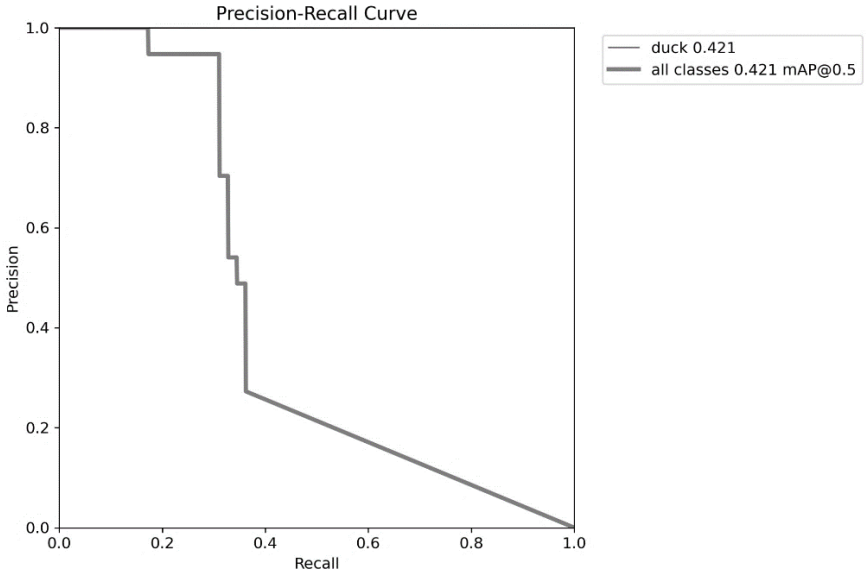


Рис. 12: Крива PR для моделі "white" на змішаному датасеті
Figure 12: PR curve for the "white" model on a mixed dataset

Для наочного порівняння продуктивності обох моделей на змішаному тестовому датасеті, ключові метрики зведено в Таблицю 2.

Таблиця 2. Порівняння продуктивності моделей на змішаному тестовому датасеті
Table 2. The comparison of model performance on the mixed test dataset

Метрика	Модель "Textured"	Модель "White"	Перевага "Textured" моделі (%)
mAP@0.5	0.682	0.421	+61.9%
mAP@0.5-0.95	0.612	0.359	+70.5%
Precision	0.965	0.855	+12.9%
Recall	0.534	0.310	+72.3%

Представлені дані чітко демонструють, що навіть за умов незначного зсуву домену модель, навчена з використанням фотореалістичної текстури, показує суттєву перевагу за всіма ключовими показниками, особливо за здатністю виявляти об'єкти (recall) та загальною якістю детекції (mAP).

3.3. Порівняльна оцінка на реальному тестовому датасеті

Ключовим етапом дослідження була валідація моделей на тестовому наборі, що складався виключно з реальних фотографій. Цей експеримент був розроблений для оцінки продуктивності в умовах повного розриву між доменами (cross-domain), що дозволяє визначити реальну практичну цінність кожного з підходів до генерації синтетичних даних.

При тестуванні на реальних даних продуктивність моделі, навченої на текстурованих зображеннях, очікувано значно знизилася порівняно з попередніми тестами. Це падіння є прямим наслідком "domain gap". Тим не менш, модель зберегла певну здатність до детектування об'єктів. Показник **mAP@0.5 склав 0.059**, а повнота (recall) – **0.123**. Це означає, що модель все ще змогла правильно ідентифікувати приблизно 12% цільових об'єктів у абсолютно новому для неї візуальному домені.

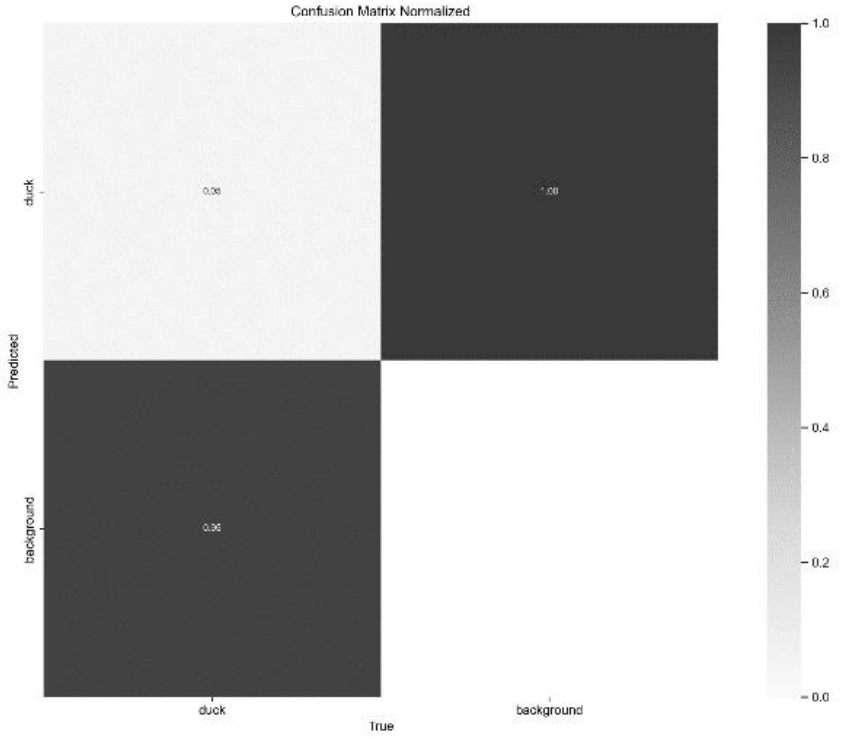


Рис. 13: Нормалізована матриця сплутування для моделі "textured" на реальному датасеті
Figure 13: Normalized confusion matrix for the "textured" model on a real dataset

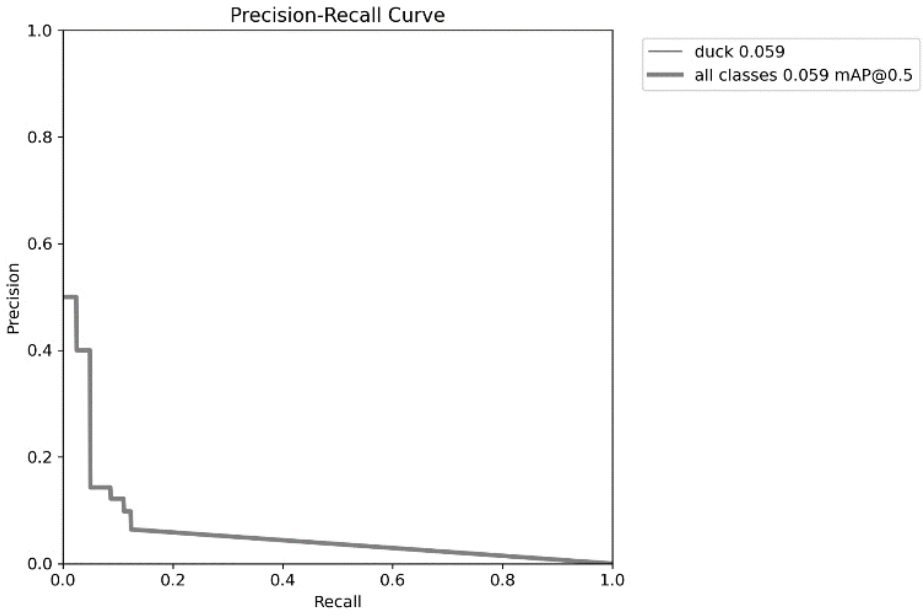


Рис. 14: Крива PR для моделі "textured" на реальному датасеті
Figure 14: PR curve for the "textured" model on a real dataset

Модель, навчена на даних без текстури, продемонструвала майже повну нездатність переносити знання на реальні зображення. Її продуктивність виявилася на порядок нижчою, ніж у текстурованої моделі, а ключові метрики наблизилися до нуля. Показник **mAP@0.5 склав лише 0.005**, а повнота (recall) – **0.012**. Фактично, модель змогла знайти лише близько 1% цільових об'єктів, що свідчить про те, що знання про суто геометричну форму об'єкта виявилися майже марними при переході до реального світу з його різноманіттям кольорів, освітлення та фонів.

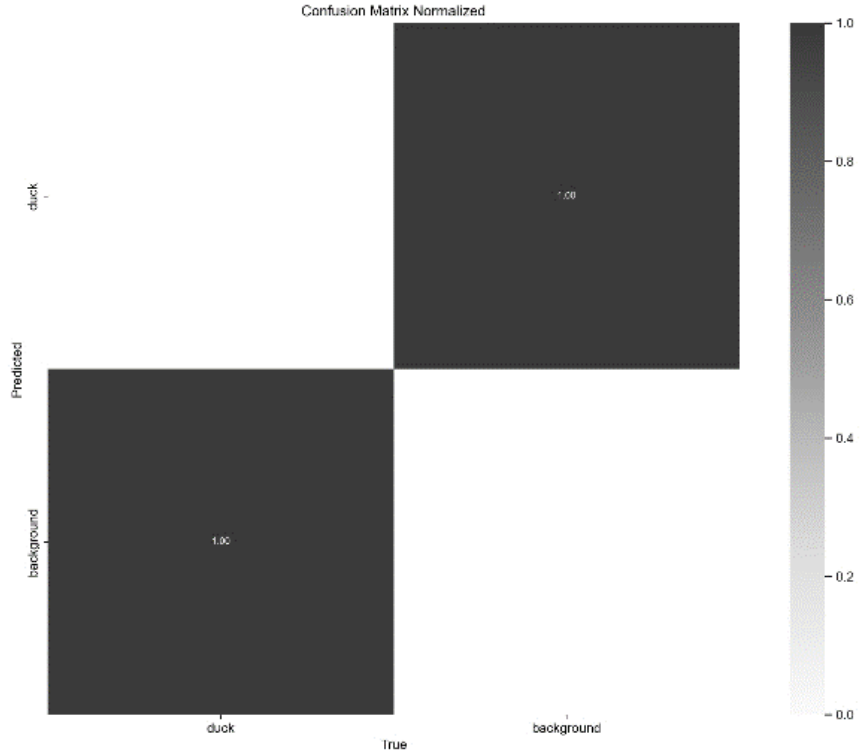


Рис. 15: Нормалізована матриця сплутування для моделі "white" на реальному датасеті
Figure 15: Normalized confusion matrix for the "white" model on a real dataset

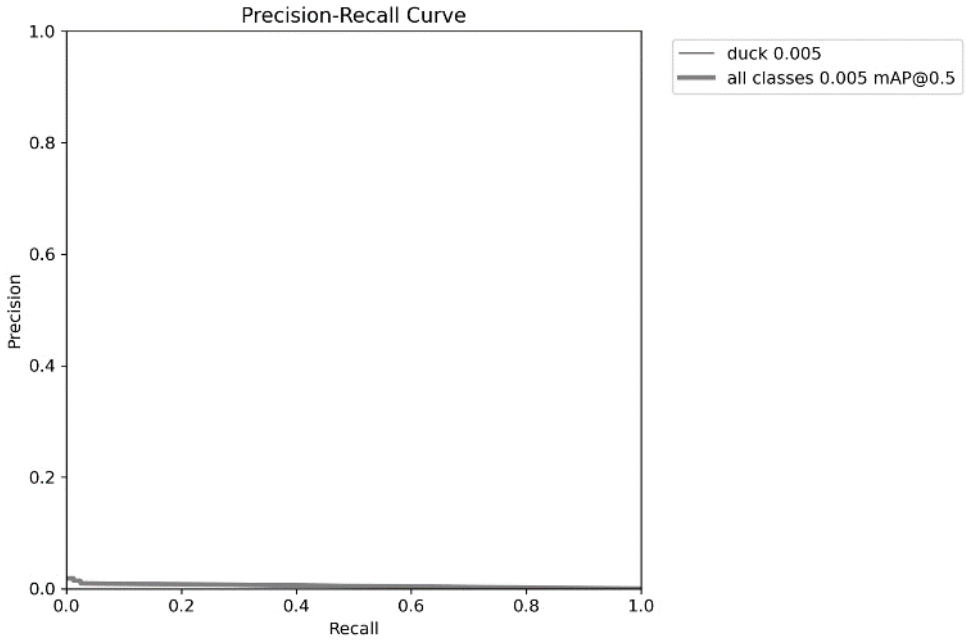


Рис. 16: Крива PR для моделі "white" на реальному датасеті
Figure 16: PR curve for the "white" model on a real dataset

Для максимально наочного порівняння ефективності обох моделей в умовах реального застосування, їхні ключові метрики зведено в Таблицю 3.

Таблиця 3. Порівняння продуктивності моделей на реальному тестовому датасеті
Table 3. The comparison of model performance on the real-world test dataset

Метрика	Модель "Textured"	Модель "White"	Перевага "Textured" моделі
mAP@0.5	0.058	0.005	Показник вищий у 11.6 разів
mAP@0.5-0.95	0.018	0.001	Показник вищий у 18.0 разів
Precision	0.064	0.015	Показник вищий у 4.3 рази
Recall	0.123	0.012	Показник вищий у 10.3 рази

Отримані дані однозначно свідчать про те, що за умов повного розриву між доменами модель, яка навчалася на синтетичних даних з фотореалістичною текстурою, демонструє на порядок вищу продуктивність. У той час як ефективність моделі, що покладалася лише на силует, виявилася близькою до нуля. Ці кількісні результати слугують емпіричною основою для подальшого аналізу та обговорення.

4. Висновок

У межах цього дослідження було проведено кількісну оцінку впливу фотореалістичної текстури цільових об’єктів у синтетичних датасетах на ефективність детектування моделями архітектури YOLOv11s при переході від симуляції до реальності (Sim2Real).

Результати експерименту повністю підтвердили висунуту гіпотезу: фотореалістична текстура є не декоративним елементом, а критично важливою структурною складовою, що забезпечує переносимість ознак та подолання «розриву між доменами» (domain gap).

Хоча обидві моделі («Textured» та «White») продемонстрували майже ідеальну та ідентичну точність на синтетичних даних (mAP@0.5 ≈ 0.995), на реальних фотографіях модель, навчена на текстурованих зображеннях, продемонструвала у 11.6 разів вищий mAP@0.5 та у 10.3 рази вищий показник повноти (recall). Це доводить, що знання про суто геометричну форму об'єкта є недостатніми для розпізнавання у реальному середовищі.

Отримані результати узгоджуються з попередніми дослідженнями робастності інтелектуальних систем у прикладних задачах [24], де показано, що стійкість моделей до зсуву вхідних розподілів є критичною для їх практичного застосування.

Висока ефективність текстурованого підходу пояснюється здатністю ранніх шарів згорткової нейронної мережі формувати універсальні низькорівневі ознаки (градієнти, кольорові переходи, мікропатерни), які є інваріантними до зміни доменів. Відсутність таких ознак у моделі без текстури призводить до її «візуальної сліпоти» на реальних об’єктах.

Найважливішим практичним висновком є те, що якісне текстурування 3D-моделей слід розглядати як стратегічний пріоритет процесу генерації даних, а не як допоміжний етап візуалізації. Інвестиції в реалістичність текстур на мікрорівні прямо конвертуються у стабільність та надійність систем комп’ютерного зору в промислових умовах.

На відміну від хаотичної доменної рандомізації (DR), метод контрольованого фотореалізму забезпечує баланс між варіативністю та фізичною правдоподібністю, створюючи синтетичні сцени з природною статистикою, що зменшує потребу в додатковій адаптації до реальних даних.

Перспективним напрямом подальших робіт є поєднання фотореалістичного текстурування з методами доменної адаптації та рандомізації матеріалів. Це дозволить забезпечити оптимальну робастність моделей для широкого спектра практичних застосувань – від автономних транспортних засобів до робототехнічних систем спостереження.

СПИСОК ЛІТЕРАТУРИ

1. Man, K.; Chahl, J. A Review of Synthetic Image Data and Its Use in Computer Vision. *J. Imaging* 2022, 8, 310.
2. Mumuni, A.; Mumuni, F. A Survey of Synthetic Data Augmentation Methods in Computer Vision. *arXiv preprint arXiv:2403.10075*, 2024.
3. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World. *In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 23-30).
4. Jackson, D., Gokhale, V., & Wyatt, J. L. (2019). Quantifying the Use of Domain Randomization for Object Localization. *arXiv preprint arXiv:1910.03438*.
5. Csurka, G. (2017). Domain Adaptation for Visual Applications: A Comprehensive Survey. *arXiv preprint arXiv:1702.05374*.
6. Wang, M., & Deng, W. (2018). Deep Visual Domain Adaptation: A Survey. *Neurocomputing*, 312, 135-153.
7. Hinterstoisser, S., Pauly, O., Heibel, H., Marek, M., & Bokeloh, M. (2019). An Annotation Saved is an Annotation Earned: Using Fully Synthetic Training for Object Instance Detection. *In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 9779-9789).
8. Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks?. *Advances in neural information processing systems*, 27.
9. Borkman, S., et al. (2021). Unity Perception: Generate Synthetic Data for Computer Vision. *arXiv preprint arXiv:2107.04259*.
10. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 779-788).
11. Koirala, A., et al. (2021). Crossing the Reality Gap: A Survey on Sim-to-Real Transferability of Robot Controllers in Reinforcement Learning. *Journal of Intelligent & Robotic Systems*, 103(4), 67.
12. Truong, J., Chernova, S., & Batra, D. (2021). Bi-directional Domain Adaptation for Sim2Real Transfer of Embodied Navigation Agents. *IEEE Robotics and Automation Letters (RA-L)*, 6(2), 2634–2641.
13. Kadian, A., Chhabra, T., Gupta, K., & Kumar, S. (2023). A Survey of Sim-to-Real Methods in RL: Progress, Prospects, and Challenges with Foundation Models. *arXiv preprint arXiv:2302.09337*.
14. Hashemifar, S., et al. (2024). Recent Advances in Deep Learning for Protein-Protein Interaction: A Review. *International Journal of Molecular Sciences*, 25(11), 5949.
15. Awais, M., et al. (2023). Don't freeze: Finetune encoders for better Self-Supervised HAR. *In Proceedings of the 2023 ACM International Symposium on Wearable Computers*.
16. Finlayson, G. D., et al. (2023). Impact of Exposure and Illumination on Texture Classification Based on Raw Spectral Filter Array Images. *Sensors*, 23(12), 5649.
17. Chung, E., et al. (2023). Inclusive Portrait Lighting Estimation Model Leveraging Graphic-Based Synthetic Data. *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*.
18. Nikolenko, S. I. (2021). *Synthetic Data for Deep Learning*. Springer Nature.
19. Picard, R. W. (2021). The Reproducibility Crisis in ML/AI: An Overview. *IEEE Open Journal of Signal Processing*, 2, 407–414.
20. Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*.
21. Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, J. (2020). Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 34(07), 12993-13000.
22. Loshchilov, I., & Hutter, F. (2016). SGDR: Stochastic Gradient Descent with Warm Restarts. *arXiv preprint arXiv:1608.03983*.

23. Goyal, P., Dollár, P., Girshick, R., Noordhuis, P., Wesolowski, L., Kyrola, A., ... & He, K. (2017). Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour. *arXiv preprint arXiv:1706.02677*.
24. Uzlov, D., Strukov, V., Hudilin, V., & Vlasov, O. (2023). Problematic issues of machine learning technology in law enforcement. Computer Science and Cybersecurity, 2, 6-15. URL:https://doi.org/10.26565/2519-2310-2023-2- 01

Korshenko Vladyslav *PhD student at the Department of Cybersecurity of Information Systems, Networks and Technologies, senior lecturer of Department of Mathematical Modeling and Data Analysis, V. N. Karazin Kharkiv National University, Svobody Square, 4, Kharkiv, Ukraine, 61077*
e-mail: v.korshenko@karazin.ua
https://orcid.org/0000-0003-2197-072X

Uzlov Dmytro *Candidate of Technical Sciences, Director of Educational and Scientific Institute of Computer Science and Artificial Intelligence, V.N. Karazin Kharkiv National University, Svobody Square, 4, Kharkiv, Ukraine, 61077*
e-mail: dmytro.uzlov@karazin.ua*;*
https://orcid.org/0000-0003-3308-424X

Assessment of the impact of photorealistic textures on the accuracy of computer vision models using synthetic datasets

Relevance. The current development of computer vision faces the problem of high cost and labor intensity of collecting real annotated data. The use of synthetic data generated in graphics engines is an effective alternative, but the main obstacle remains the "domain gap," which reduces the accuracy of models on real images.
The goal of this work is to quantitatively assess the impact of the photorealistic texture of the target object on the detection efficiency of YOLO models when transitioning from simulation to reality (Sim2Real).
The research methodology is based on a controlled experiment in the Unity environment, where two identical synthetic datasets were generated, differing only in the type of 3D model texture: highly detailed photorealistic ("Textured") and monochrome white ("White"). The models were trained based on the YOLOv11s architecture using a transfer learning strategy and a two-step fine-tuning process. The results were validated on an independent set of exclusively real photographs.
Results. Both models, trained on two datasets ("Textured" and "White"), achieved almost identical accuracy on synthetic validation data (mAP@0.5 ≈ 0.995). However, on real photos, the "Textured" model demonstrated 11.6 times higher mAP@0.5 compared to the "White" model. The recall for the textured model was 10.3 times higher than for the model that relied solely on geometric shape.
Conclusions. Photorealistic texture is a critical factor for successful Sim2Real transfer. It ensures the formation of universal low-level features in the early layers of the neural network, which are necessary for recognizing objects in a real environment. High-quality texturing of 3D assets should be considered a strategic priority rather than an auxiliary stage of visualization.
***Keywords**: synthetic data, computer vision, object detection, domain gap, model robustness, domain shift resilience.*

Надійшла у першій редакції 11.01.2026; рекомендована до друку 12.03.2026; опублікована 03.04.2026.
The first version has been received on 11.01.2026; was recommended for printing on 12.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-05

UDC 004.738.5:004.9-055.52

Savchenko Mykhailo *student*
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37 Beresteiskyi Ave., Kyiv-56, Ukraine, 03056
https://orcid.org/0009-0005-9441-3467
e-mail: its30316@gmail.com

Sulima Svitlana *PhD, assistant professor;*
assistant professor of the Department of Information technologies in telecommunications, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37 Beresteiskyi Ave., Kyiv-56, Ukraine, 03056
e-mail: itssulima@gmail.com;
https://orcid.org/0000-0002-6333-7693

Modular JavaScript library for ensuring web interface accessibility in accordance with WCAG 2.2

Relevance. Web accessibility has become a critical aspect of modern web development, considering the needs of more than 1.3 billion people with disabilities worldwide. Despite the existence of WCAG standards, the vast majority of websites remain inaccessible, highlighting the demand for comprehensive yet easy-to-integrate tools that address key accessibility challenges.
Purpose. The main goal is to develop a modular JavaScript library that provides comprehensive web interface accessibility enhancements in accordance with WCAG 2.2, while maintaining simplicity of integration and high performance.
Research Methods. The research applied a user-centered iterative development methodology with step-by-step validation of features through scripted evaluation, comparative testing with existing solutions, and the implementation of a browser extension for practical verification.
Results. A modular JavaScript library was developed consisting of seven independent components (dark mode, high contrast, keyboard navigation, text scaling, focus enhancement, dyslexia support, double-click protection), each addressing specific WCAG 2.2 success criteria. The effectiveness of the components was demonstrated through measurable improvements: enhanced contrast ratios (from 3.8:1 to 21:1), a 25% reduction in keystrokes for navigation, increased focus visibility (contrast improvement from 1 to 6.5), and full compliance with dyslexia readability parameters. Real-time interaction and dynamic content adaptation further improve user experience.
Conclusions. The proposed solution bridges the gap between fragmented accessibility tools by offering a unified approach with a high level of modularity. The library has demonstrated practical feasibility through a browser extension and is ready for integration into existing web projects. The proposed architecture provides a robust foundation for future research and development in the field of digital accessibility.
Keywords: *web accessibility, WCAG 2.2, JavaScript library, browser extension, interface adaptation, support for users with disabilities, inclusivity.*

Як цитувати: Savchenko M. Sulima S. Modular JavaScript library for ensuring web interface accessibility in accordance with WCAG 2.2. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.59-72. https://doi.org/10.26565/2304-6201-2026-69-05

How to quote: M. Savchenko S. Sulima,“ Modular JavaScript library for ensuring web interface accessibility in accordance with WCAG 2.2”, *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 69, pp. 59-72, 2026. hhttps://doi.org/10.26565/2304-6201-2026-69-05

Introduction

The publication of the Web Content Accessibility Guidelines (WCAG) introduced a formalized approach to web accessibility. With each iteration, the guidelines emphasized the need for developers to create accessible content and applications, paving the way for the development of specialized JavaScript libraries tailored for compliance with these standards [1]. The introduction of WCAG 2.2 further refined accessibility criteria, leading to the creation of modular JavaScript libraries that not only facilitate compliance but also enhance the overall user experience [1]. These libraries focus on reusability and modularity, allowing developers to easily integrate accessibility features into their projects while adhering to the latest guidelines. Consequently, the evolution of these libraries has become integral to modern web development practices, ensuring inclusivity for all users [2][3]. As the web continues to evolve, the

emphasis on accessibility remains critical, prompting ongoing innovation in the creation of tools and libraries that align with the latest standards and foster an inclusive digital environment for diverse user populations.

Web accessibility has become a fundamental requirement in modern web development, with over 1.3 billion people worldwide living with some form of disability according to the World Health Organization. Despite the existence of accessibility standards such as the Web Content Accessibility Guidelines [4] (WCAG), many websites remain inaccessible to users with disabilities. Recent studies indicate that 98% of websites have at least one WCAG failure, highlighting the urgent need for comprehensive accessibility solutions.

The challenge lies not only in awareness but also in the complexity of implementing accessibility features across diverse web interfaces. Traditional approaches often require extensive manual coding, specialized expertise, and significant development time, creating barriers to widespread adoption. Furthermore, existing solutions are frequently fragmented, addressing only specific accessibility needs rather than providing comprehensive support.

This research addresses these challenges by developing a modular JavaScript library that provides comprehensive accessibility enhancements while maintaining simplicity of integration and use. The library's modular design allows developers to implement specific accessibility features as needed, reducing complexity while ensuring compliance with international standards.

1. Research Objectives

The primary objectives of this research are:

- to develop a comprehensive, modular JavaScript library for web accessibility enhancement
- to ensure compliance with WCAG 2.2 standards across all implemented features
- to create an intuitive integration process that reduces implementation barriers
- to validate the library's effectiveness through practical testing and measurement
- to demonstrate the solution's applicability through a browser extension implementation

The scientific novelty of this work lies in several key areas:

1. Modular Architecture Innovation: Unlike existing accessibility solutions that often provide monolithic implementations, this research introduces a truly modular approach where each accessibility feature operates independently while maintaining seamless integration capabilities. This architecture allows for selective implementation based on specific user needs and project requirements.
2. Comprehensive WCAG 2.2 Mapping: The research provides a systematic mapping of each library component to specific WCAG 2.2 criteria, establishing a clear framework for compliance verification and effectiveness measurement.
3. Dynamic Adaptation System: The library implements dynamic adaptation mechanisms that respond to user interactions and preferences in real-time, providing personalized accessibility experiences without requiring page reloads or complex configuration.
4. Quantitative Effectiveness Measurement: The research introduces novel metrics and testing methodologies for quantifying accessibility improvements, including contrast enhancement ratios, navigation efficiency measurements, and focus visibility calculations.
5. Browser Extension Validation Framework: The development of a companion browser extension serves as both a practical implementation example and a validation tool, demonstrating real-world applicability and effectiveness.

2. Current State of Web Accessibility

Web accessibility research has evolved significantly since the introduction of the first WCAG guidelines in 1999. Current accessibility solutions can be categorized into several approaches: server-side implementations, client-side JavaScript libraries, browser extensions, and assistive technologies.

Server-side solutions, while comprehensive, require significant infrastructure changes and may not be feasible for all organizations. Client-side JavaScript libraries offer more flexibility but often lack comprehensive coverage of accessibility needs. Browser extensions provide user-controlled accessibility enhancements but typically operate independently of website design considerations.

Several JavaScript libraries address specific accessibility concerns. Libraries like "ally.js"[5] focus on focus management, while "a11y-dialog"[6] specializes in accessible modal dialogs. However, these solutions typically address single aspects of accessibility rather than providing comprehensive coverage.

The gap in existing solutions lies in the absence of a unified, modular approach that combines multiple accessibility features while maintaining independence between components. This research fills this gap by providing a comprehensive solution with a modular architecture.

WCAG 2.2 introduces additional success criteria that address mobile accessibility, cognitive disabilities, and low vision requirements. Implementing these standards requires detailed understanding of user needs and technical implementation strategies. This research addresses these challenges by providing pre-built, tested components that ensure compliance.

3. Development Approach.

The development methodology follows a user-centered design approach combined with iterative development and continuous testing. The process consists of several phases:

1. Requirements Analysis: Comprehensive analysis of WCAG 2.2 requirements and user needs.
2. Component Design: Modular architecture design ensuring independence and interoperability.
3. Implementation: Development of individual components with focus on performance and compatibility.
4. Testing: Automated and manual testing using industry-standard tools.
5. Validation: Browser extension development for real-world testing.

The library architecture is based on several key principles:

1. Modularity: Each component operates independently, allowing selective implementation and reducing code bloat.
2. Interoperability: Components can work together seamlessly when multiple accessibility features are needed.
3. Performance: Minimal impact on page load times and runtime performance through efficient code design and lazy loading.
4. Compatibility: Cross-browser compatibility ensuring functionality across modern web browsers.
5. Extensibility: Clear interfaces for adding new components and extending existing functionality.

4. Library Architecture and Implementation

The library employs a modular architecture where components are organized in a dedicated directory structure. The main architecture consists of:

```
accessibility-library/
├── components/
│   ├── highContrast.js
│   ├── focusMode.js
│   ├── darkMode.js
│   ├── keyboardNav.js
│   ├── doubleClick.js
│   ├── dyslexiaMode.js
│   └── increaseText.js
├── core/
│   └── loader.js
└── utils/
    └── helpers.js
```

Each component is designed with modularity in mind, enabling separate use. A folder structure supports extensibility, making it easy to add new components without making major changes to the existing code.

This clear project structure simplifies developers' work and provides a convenient entry mechanism for new team members. The architecture facilitates rapid deployment and engagement, making the library appealing to developers looking to swiftly and effectively enhance the user experience of their web applications.

4.1. Component Specifications

High Contrast Component (highContrast.js) addresses WCAG 2.2 criteria 1.4.3 (Contrast Minimum), 1.4.6 (Contrast Enhanced), and 1.4.1 (Use of Color). It implements dynamic style injection to achieve maximum contrast ratios between text and background elements.

Implementation features:

- Dynamic CSS injection for high contrast styles
- Preservation of original styles for reversibility
- Support for complex layouts and nested elements
- Real-time application without page refresh

Technical approach: the component creates a dedicated <style> element with high contrast CSS rules that override existing styles while maintaining layout integrity. The implementation uses CSS specificity and !important declarations strategically to ensure consistent application across diverse website designs.

Focus Enhancement Component (focusMode.js) enhances visual focus indicators to meet WCAG 2.2 criteria 1.4.13 (Content on Hover or Focus) and 2.4.7 (Focus Visible). It provides clear visual feedback for keyboard navigation and improves usability for users with motor impairments.

Implementation features:

- Enhanced focus outlines with customizable colors
- Shadow effects for improved visibility
- Support for all focusable elements
- Dynamic event handling for hover and focus states

Dark Mode Component (darkMode.js) addresses WCAG criteria 1.4.3, 1.4.11 (Non-text Contrast), and provides relief for users with light sensitivity or certain neurological conditions.

Implementation features:

- Intelligent color inversion algorithms
- Preservation of image and media content
- Site-specific optimizations for popular platforms
- Selective element targeting to maintain design coherence

Keyboard Navigation Component (keyboardNav.js) enhances keyboard navigation capabilities beyond standard browser implementations, addressing WCAG criteria 2.1.1 (Keyboard) and 2.1.2 (No Keyboard Trap).

Implementation features:

- Arrow key navigation between focusable elements
- Spatial awareness for logical navigation flow
- Visual feedback for current focus position
- Skip navigation for complex layouts

Novel algorithm: the component implements a spatial navigation algorithm that calculates element positions and determines the most logical navigation path based on geometric relationships rather than DOM order.

Double-Click Protection Component (doubleClick.js) provides protection against accidental activations, particularly beneficial for users with motor impairments or tremors. It addresses WCAG criteria 3.3.2 (Labels or Instructions) and 2.5.1 (Pointer Gestures).

Implementation features:

- Configurable delay between clicks
- Element-specific handling for different interaction types
- Visual feedback during delay periods
- Accessibility announcements for screen readers

Dyslexia Support Component (dyslexiaMode.js) implements typography and spacing modifications to improve readability for users with dyslexia, addressing WCAG criteria 1.4.5 (Images of Text) and 3.1.5 (Reading Level).

Implementation features:

- OpenDyslexic font integration
- Enhanced letter and word spacing
- Improved line height and paragraph spacing
- ARIA live announcements for mode changes

Text Scaling Component (increaseText.js) provides dynamic text size adjustment while maintaining layout integrity, addressing WCAG criteria 1.4.4 (Resize Text) and 1.4.10 (Reflow).

Implementation features:

- Proportional scaling across all text elements
- Layout preservation during scaling
- Original size restoration capability
- Accessibility announcements for changes

4.2. Integration and Usage

The library provides multiple integration methods to accommodate different development workflows (Fig. 1).

Direct Integration:

```javascript
// Load specific components
import { enableHighContrast } from './components/highContrast.js';
import { enableDarkMode } from './components/darkMode.js';

// Activate features
enableHighContrast(true);
enableDarkMode(true);
```

Configuration-Based Integration:

```javascript
// Configure multiple features
const accessibilityConfig = {
  highContrast: true,
  darkMode: false,
  textScale: 1.2,
  keyboardNav: true
};

AccessibilityLibrary.init(accessibilityConfig);
```

Fig. 1 Integration
Рис. 1 Інтеграція

4.3. Browser Extension Implementation

To validate the library's practical applicability, a browser extension was developed that demonstrates real-world usage scenarios.

The extension architecture includes following core files (Fig. 2):

- manifest.json: extension configuration and permissions
- popup.html: user interface template
- popup.js: user interaction handling
- content.js: content script for DOM manipulation
- styles.css: interface styling
- lib/components/ directory: this directory contains components that are responsible for specific accessibility features. Each component is a separate JavaScript file that implements a specific accessibility feature on web pages.

The extension dynamically loads library components based on user selections, demonstrating the modular architecture's flexibility and performance benefits.

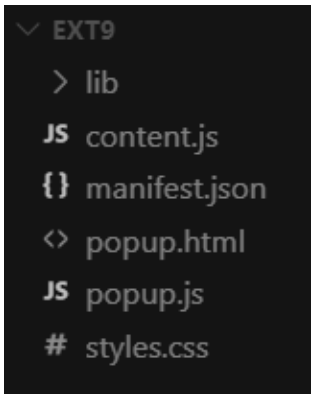


Fig. 2 Extension structure
Рис. 2 Структура розширення

This structure allows for flexibility and ease of expansion, as each individual part of the extension performs a clearly defined role, and centralized management of components and settings allows the interface to be quickly adapted to user needs.

The extension uses a multi-level architecture, where the main component is the content script (content.js). This script is automatically loaded on every page the user visits and acts as a coordinator, linking the user interface with accessibility features.

The accessibility components, located in the lib/components/ directory, are standalone modules responsible for specific functions (e.g., dark mode or text enlargement). They are loaded dynamically depending on the user's choice, which optimizes performance. When the user activates a specific feature via the popup interface, content.js receives a signal and initiates the loading of the corresponding component. After loading, the component makes changes to the DOM of the web page to improve accessibility without the need to reload the page.

In addition, the system saves user settings, and these settings are automatically restored the next time the page is visited. This architecture ensures high performance and allows you to easily add new features by creating new modules in the lib/components/ directory and registering them in the system.

The interface is implemented via a popup window (popup.html and popup.js) – Fig. 3.

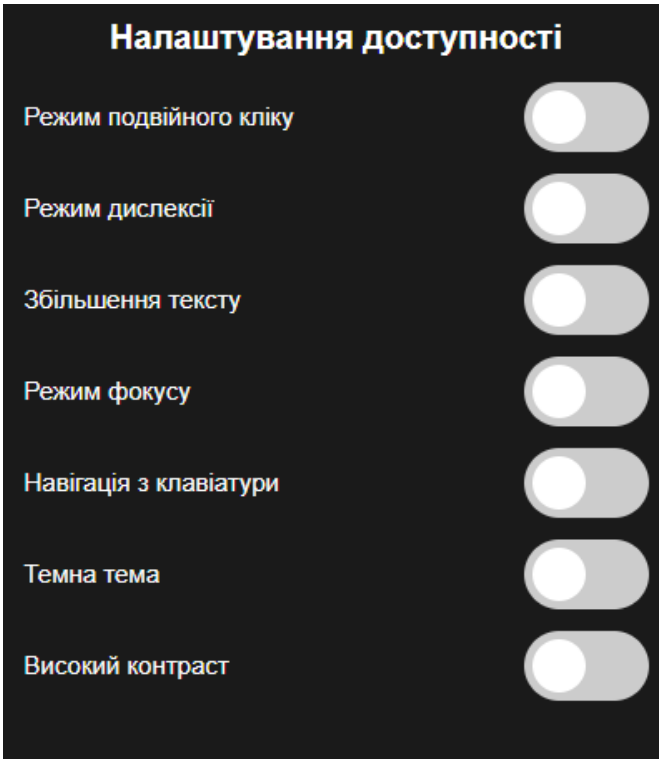


Fig. 3 User interface of the extension
Рис. 3 Користувацький інтерфейс розширення

4.4. Browser Extension Implementation

Let's take a few functions as an example and evaluate their effectiveness and consistency with the purpose.

First, we can evaluate the effectiveness of the high contrast mode function.

Using the WAVE Evaluation Tool, we will visit a website, for example, KPI Campus:

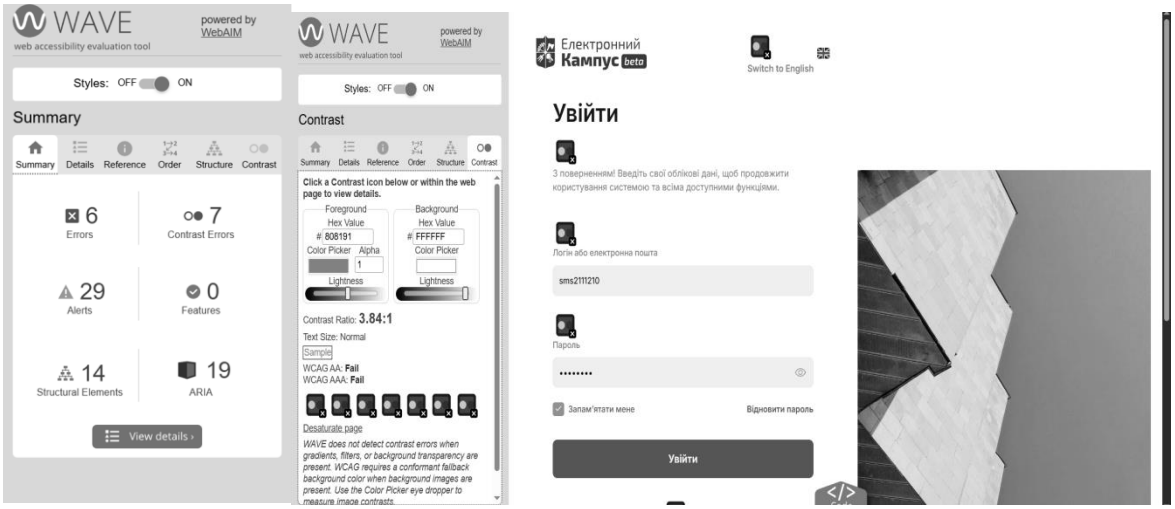


Fig. 4 Contrast evaluation with the WAVE tool without using the extension

Рис. 4 Оцінка контрасту за допомогою інструменту WAVE без використання розширення

As we can see, there are elements (7) on the website that do not meet the WCAG contrast standards and have a value of ~3.8:1, which is less than the standard.

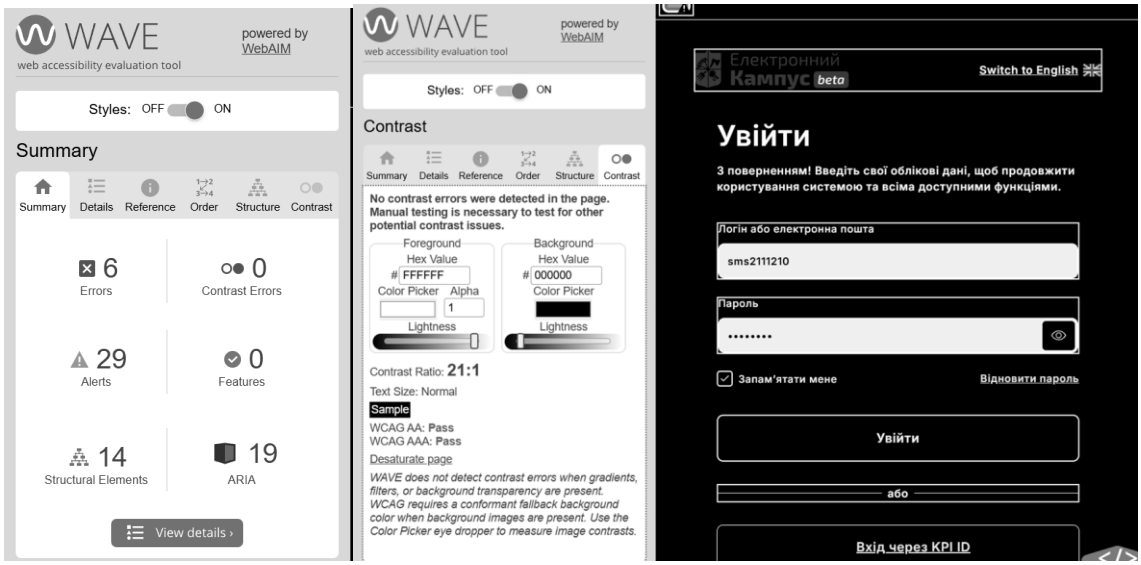


Fig. 5 Contrast evaluation with the WAVE tool without using the extension

Рис. 5 Оцінка контрасту за допомогою інструменту WAVE без використання розширення

Here is the path you need to follow from start to finish with standard keyboard navigation – Fig. 6.

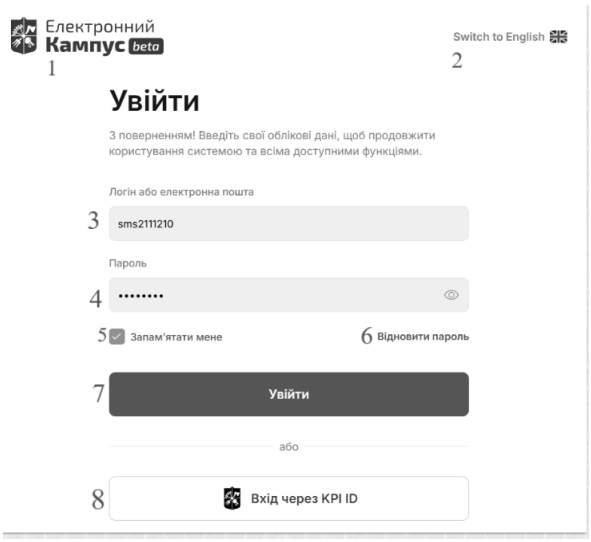


Fig. 6 Path from start to finish with standard navigation
Рис. 6 Шлях від початку до кінця зі стандартною навігацією

And here is the path with the developed navigation – Fig. 7.

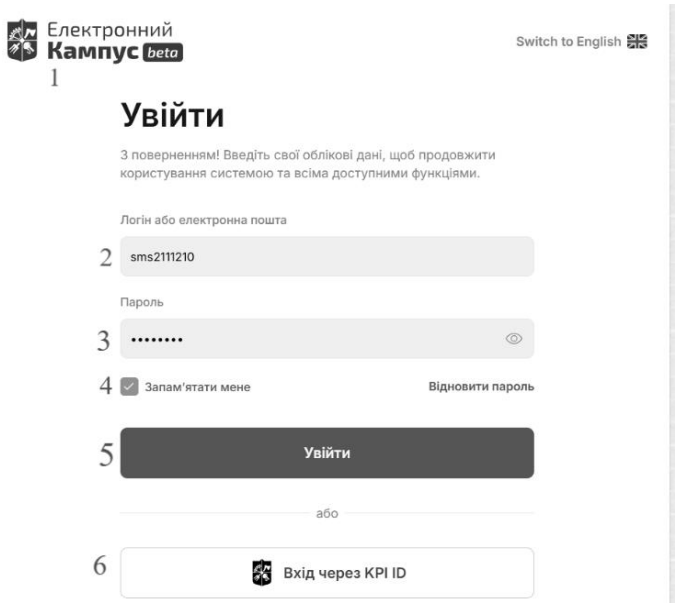


Fig. 7 Path from start to finish using extension navigation
Рис. 7 Шлях від початку до кінця за допомогою навігації з розширенням

As we can see, even in such a simple example, where there are not many elements and they are all quite consistent, in the first case you need to go through 8 elements, in the second 6. This means that the number of actions is reduced by 25%, which significantly increases the speed and convenience of navigation. And this is the minimum value, because on more complex web interfaces with a larger number of elements, the result will be much longer for standard navigation, which cannot be said about navigation with this extension.

To evaluate the focus mode and its compliance with the specified goal, you can develop a small script that would check the contrast of elements. For example, the following script allows you to evaluate how noticeable the focus is on interactive elements: when you hover the cursor or focus (via the keyboard), it analyzes how the element stands out — using an outline or box shadow — and calculates the contrast between these styles and the background color. The results are displayed in the console and help verify that the focus meets accessibility requirements.

```
(() => {
  const focusableSelector = 'a, button, input, select, textarea, [tabindex]:not([tabindex="-1"])';

  function getRGB(str) {
    const match = str.match(/rgba?\((\d+),\s*(\d+),\s*(\d+)/);
    return match ? match.slice(1, 4).map(Number) : null;
  }

  function luminance(rgb) {
    if (!rgb) return 0;
    const [r, g, b] = rgb.map(v => {
      v /= 255;
      return v <= 0.03928 ? v / 12.92 : Math.pow((v + 0.055) / 1.055, 2.4);
    });
    return 0.2126 * r + 0.7152 * g + 0.0722 * b;
  }

  function contrast(rgb1, rgb2) {
    if (!rgb1 || !rgb2) return 1;
    const l1 = luminance(rgb1);
    const l2 = luminance(rgb2);
    return ((Math.max(l1, l2) + 0.05) / (Math.min(l1, l2) + 0.05)).toFixed(2);
  }

  function analyze(el) {
    const cs = getComputedStyle(el);
    const bg = getRGB(cs.backgroundColor);
    const outlineColor = cs.outlineStyle !== 'none' ? getRGB(cs.outlineColor) : null;
    const shadowMatch = cs.boxShadow.match(/rgba?\([^)]+\)/g);
    const shadowColor = shadowMatch ? getRGB(shadowMatch.at(-1)) : null;

    const contrastOutline = contrast(outlineColor, bg);
    const contrastShadow = contrast(shadowColor, bg);

    console.clear();
    console.log('🔍 Элемент:', el);
    console.log('Outline:', cs.outline);
    console.log('Box-shadow:', cs.boxShadow);
    console.log('Контраст outline:', contrastOutline);
    console.log('Контраст shadow:', contrastShadow);
  }

  function handler(e) {
    const el = e.target.closest(focusableSelector);
    if (el) analyze(el);
  }

  document.addEventListener('mousemove', handler);
  document.addEventListener('focusin', handler);
```

Fig.8 Script code for evaluating focus mode
Рис. 8 Код скрипта для визначення режиму фокусування

For example, here is the result without focus mode enabled – Fig. 9.

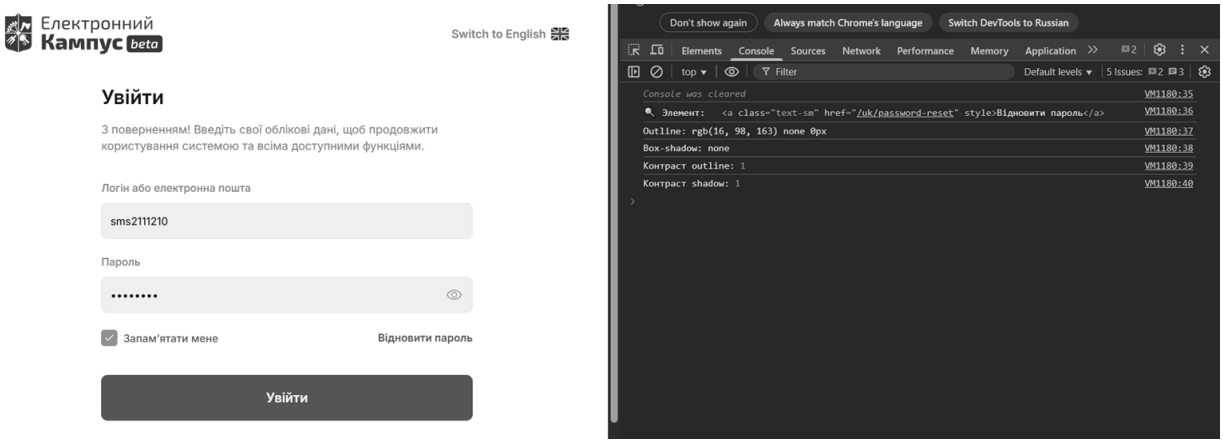


Fig. 9 Script result without focus mode enabled
Рис. 9 Результат виконання скрипта без увімкненого режиму фокусування

And here is the result with focus mode enabled - Fig. 10.

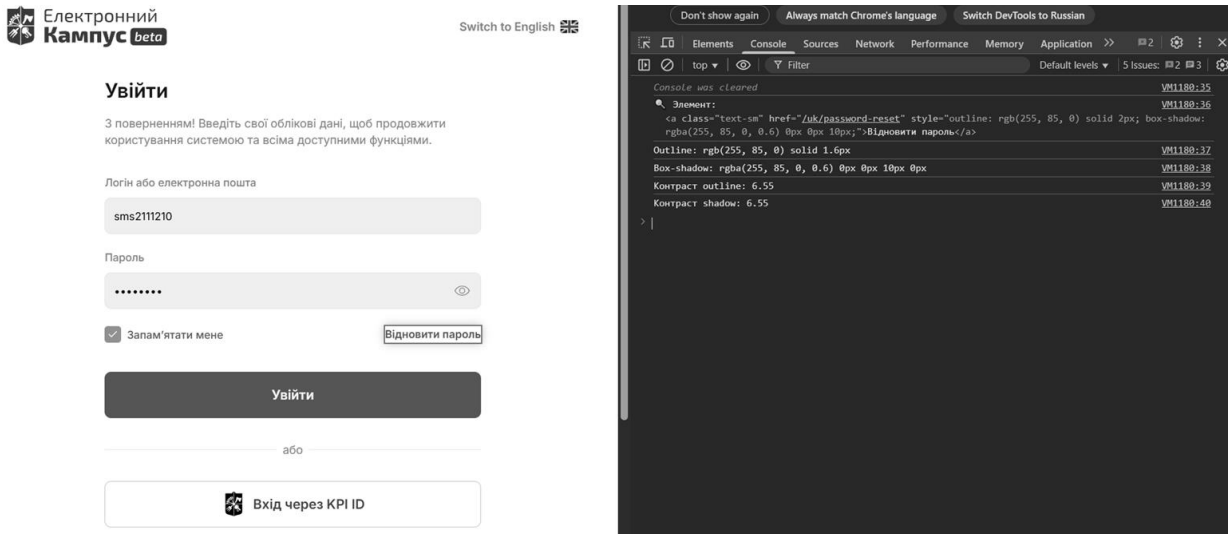


Fig. 10 Script result with focus mode enabled
Рис. 10 Результат виконання скрипта при увімкненому режимі фокусування

As we can see, without active focus mode, the element has no visible outline or shadow, and the contrast level is 1 — the focus is visually invisible. However, with the specified style (outline and box-shadow with a bright color), the contrast increases to 6.55, which is sufficient to ensure a noticeable focus on the element according to WCAG standards (minimum 3:1 for visible focus).

Next, we can evaluate the effectiveness in terms of compliance with the purpose of dyslexia mode. The script analyzes key parameters that affect readability for people with dyslexia: the font used, the spacing between letters, words, and lines, and the presence of dynamic support through the aria-live attribute. The result is a summary score that helps to quickly determine how well the text meets the basic accessibility criteria for this mode.

```
// Функція для отримання кольору у форматі RGB
function getRGB(color) {
    const div = document.createElement('div');
    div.style.backgroundColor = color;
    document.body.appendChild(div);
    const computed = window.getComputedStyle(di
    document.body.removeChild(div);
    const rgb = computed.match(/\d+/g).map(Numb
    return rgb.length >= 3 ? rgb : [255, 255, 2
}

// Функція для обчислення яскравості (залишена
function getLuminance(rgb) {
    const [r, g, b] = rgb.map(v => v / 255);
    const a = [r, g, b].map(v => {
        return v <= 0.03928 ? v / 12.92 : Math.
    });
    return 0.2126 * a[0] + 0.7152 * a[1] + 0.07
}

// Основна функція для оцінки параметрів режиму
function evaluateDyslexiaMode() {
    const paragraphs = document.querySelectorAl
    const results = {
        font: { score: 0, value: '' },
        letterSpacing: { score: 0, value: 0 },
        wordSpacing: { score: 0, value: 0 },
        lineHeight: { score: 0, value: 0 },
        ariaLive: { score: 0, value: 'Не знайде
    };

    if (paragraphs.length > 0) {
        const p = paragraphs[0];
        const style = window.getComputedStyle(p
        const fontFamily = style.fontFamily.toL
        results.font.value = fontFamily;

        // Оцінка шрифту
        if (fontFamily.includes('opendyslexic')
            results.font.score = 2;
        } else if (!fontFamily.includes('serif'
            results.font.score = 1;
        } else {
            results.font.score = 0;
        }

        // Інтервал між літерами
        const letterSpacing = parseFloat(style.
        const fontSize = parseFloat(style.fontS
        results.letterSpacing.value = letterSpa
        if (letterSpacing >= 0.12 * fontSize) {
            results.letterSpacing.score = 1;
        }

        // Інтервал між словами
        const wordSpacing = parseFloat(style.wordSpacing);
        results.wordSpacing.value = wordSpacing;
        if (wordSpacing >= 0.16 * fontSize) {
            results.wordSpacing.score = 1;
        }

        // Міжрядковий інтервал
        const lineHeight = parseFloat(style.lineHeight) / fontSize;
        results.lineHeight.value = lineHeight;
        if (lineHeight >= 1.5) {
            results.lineHeight.score = 1;
        }
    }

    // Перевірка наявності aria-live
    const ariaLiveElement = document.querySelector('[aria-live]');
    if (ariaLiveElement) {
        results.ariaLive.value = ariaLiveElement.getAttribute('aria-live');
        if (['polite', 'assertive'].includes(results.ariaLive.value)) {
            results.ariaLive.score = 1;
        }
    }

    // Виведення результатів у консоль
    console.log('📊 Оцінка параметрів для режиму дислексії:');
    console.log(1. Шрифт: ${results.font.value} → Бал: ${results.font.score});
    console.log(2. Інтервал між літерами: ${results.letterSpacing.value}px → Бал: ${results.letterSpaci
    console.log(3. Інтервал між словами: ${results.wordSpacing.value}px → Бал: ${results.wordSpacing.sc
    console.log(4. Міжрядковий інтервал: ${results.lineHeight.value} → Бал: ${results.lineHeight.score}
    console.log(5. aria-live: ${results.ariaLive.value} → Бал: ${results.ariaLive.score});

    const totalScore = Object.values(results).reduce((sum, r) => sum + r.score, 0);
    const maxScore = 6; // 2 бали за шрифт, по 1 за інші 4 критерії

    console.log(\n🔹 Загальна підсумкова оцінка: ${totalScore}/${maxScore} балів);
}

// Виклик функції
evaluateDyslexiaMode();
```

Fig. 11 Script code for evaluating dyslexia mode
Рис. 11 Код скрипта для активації режиму дислексії

Next, we evaluate the effectiveness of dyslexia mode based on key parameters such as font, spacing, and aria-live support. The script calculates the score for each criterion.

Below are examples of results from the console:

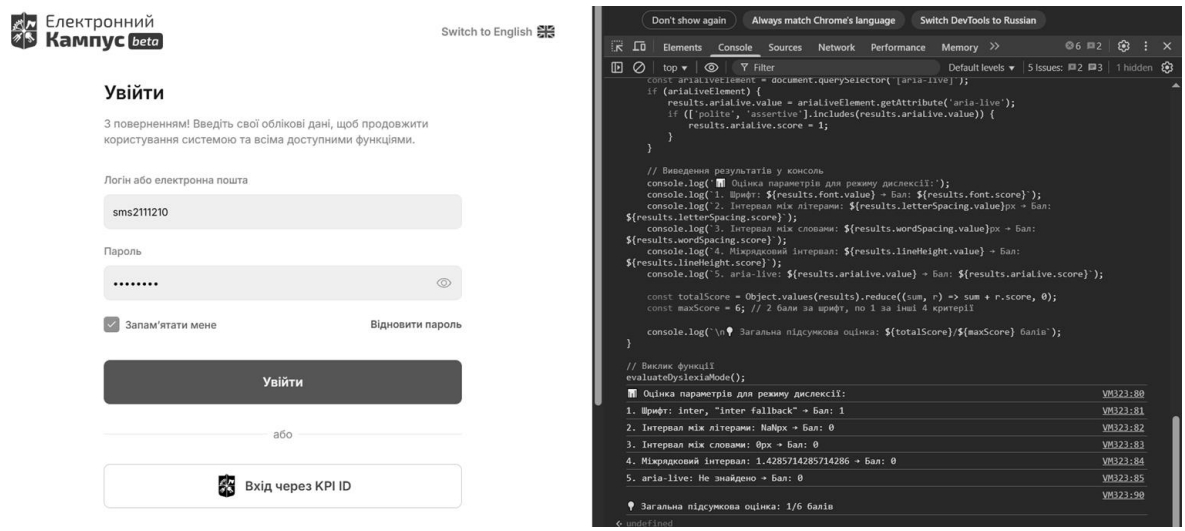


Fig. 12 Script result without dyslexia mode enabled

Рис. 12 Результат виконання скрипта без увімкненого режиму дислексії

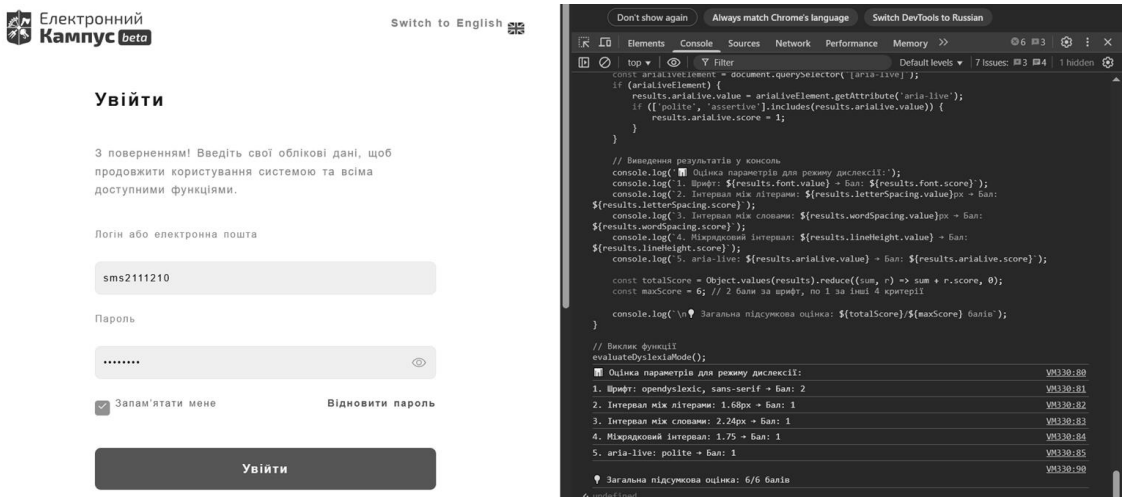


Fig. 13 Script result with dyslexia mode enabled

Рис. 13 Результат виконання скрипта з увімкненим режимом для людей з дислексією

Initially, the dyslexia mode score was low – only 1 out of 6 points, indicating that the web resource was not sufficiently inclusive. After the changes were made, the result improved to a maximum of 6 out of 6, which means full compliance with the basic requirements for comfortable reading by users with dyslexia.

Now let's check the effectiveness of dark mode. This script collects the background colors of the main blocks of the page (body, header, main, section), converts them to RGB numerical values, and calculates their brightness (lux) using a formula that takes into account the perception of color by the human eye: $0.2126 \times R + 0.7152 \times G + 0.0722 \times B$, where R, G, B are the values of the red, green, and blue channels, respectively.

The script then calculates the average brightness of the background to estimate the overall brightness of the page.

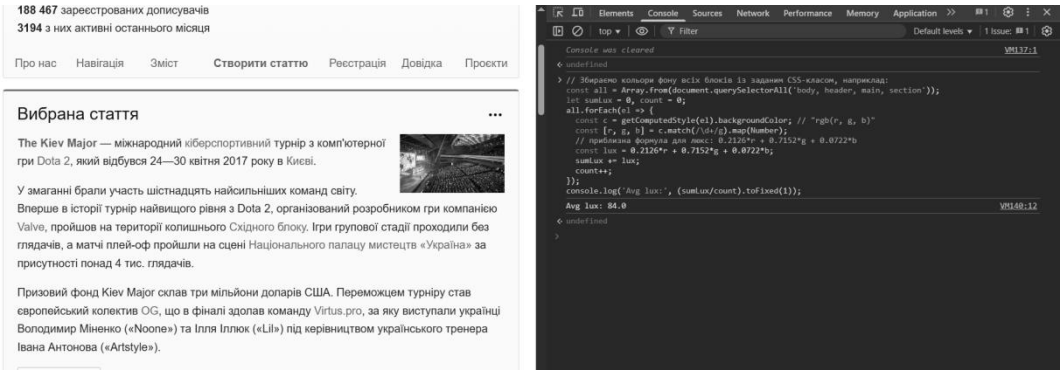


Fig. 14 Result of the script without dark mode enabled
Рис. 14 Результат виконання скрипта без увімкненого темного режиму

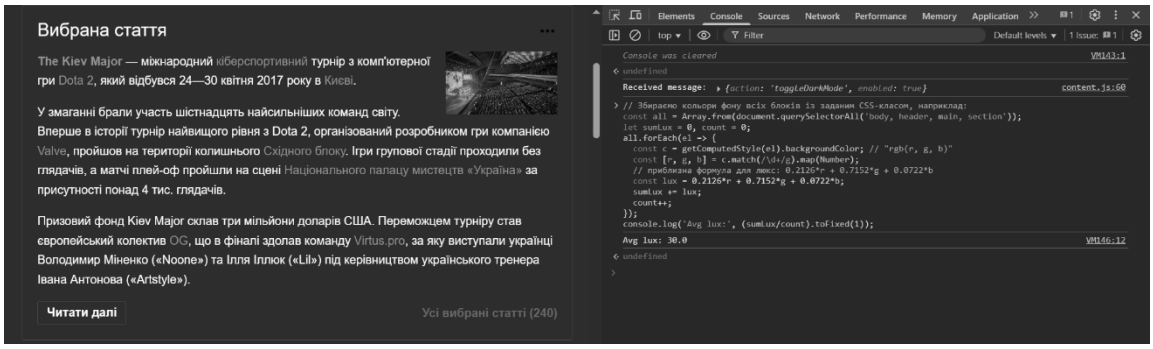


Fig. 15 Result of the script with dark mode enabled
Рис. 15 Результат виконання скрипта з увімкненим темним режимом

We tested the brightness of the main page blocks. Before enabling dark mode, the average lux value was 84.0, which corresponds to a fairly bright background. After activating dark mode, this indicator decreased to 30.0, which indicates a significant darkening of the interface.

According to WCAG recommendations and accessibility practices, optimizing brightness levels helps reduce eye strain, especially for users with photosensitivity or visual impairments. Dark mode with reduced lux levels increases viewing comfort and makes the interface more accessible to a wider range of users.

Let's move on to evaluating the effectiveness of double-click implementation. To do this, we used a script that analyzes all interactive elements on the page and calculates what percentage of them belong to the types supported by the doubleClick component. These are considered to be those that can correctly process a second click — in particular, links, buttons, checkboxes, and text blocks.

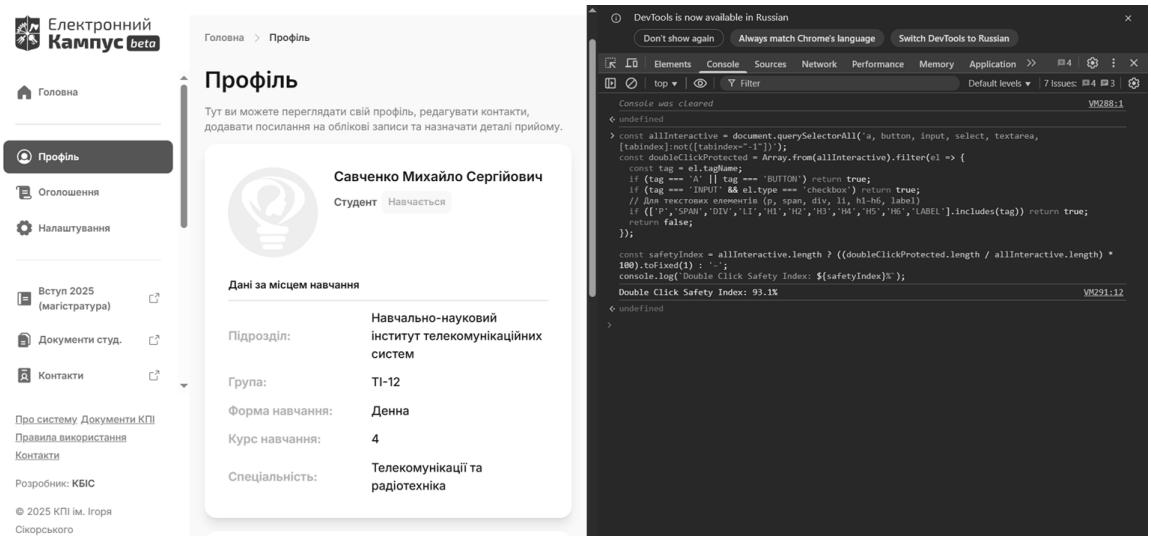


Fig. 16 The result of the script that checks the effectiveness of the double-click function
Рис. 16 Результат виконання скрипта, що перевіряє працездатність функції подвійного клацання

This means that 93.1% of all interactive elements on the page (links, buttons, checkboxes, text blocks, etc.) are protected by double-click logic. In other words, these elements have a built-in delay between clicks, which prevents accidental or premature activation. This implementation significantly reduces the risk of accidental interactions — for example, accidentally following a link or clicking a button. This is especially important for users with motor impairments, tremors, reaction delays, or those who use alternative input devices. The introduction of double-clicking as an action filter improves the accuracy of interaction with the interface, increases the level of control for the user, and contributes to a more inclusive and accessible experience.

Conclusions

This research successfully developed and validated a comprehensive, modular JavaScript library for enhancing web interface accessibility. The library addresses critical gaps in existing accessibility solutions by providing a unified, easy-to-integrate approach that maintains compliance with WCAG 2.2 standards while offering significant flexibility and performance benefits.

The scientific significance of this work extends beyond the immediate technical contributions:

1. Methodological Advancement: The research establishes new approaches for quantifying accessibility improvements and provides frameworks for systematic accessibility enhancement.
2. Standardization Contribution: The comprehensive mapping to WCAG 2.2 criteria and systematic implementation approach contributes to standardization efforts in accessibility technology.
3. Interdisciplinary Integration: The work bridges computer science, human-computer interaction, and disability studies, providing insights valuable across multiple disciplines.

The practical impact of this research is demonstrated through:

1. Immediate Applicability: The library can be immediately integrated into existing web projects with minimal modification.
2. Developer Accessibility: The modular approach reduces barriers to accessibility implementation for developers with varying expertise levels.
3. User Empowerment: The browser extension demonstrates how accessibility tools can be made directly available to users.
4. Cost Reduction: The library reduces the cost and complexity of implementing comprehensive accessibility features.

This research establishes a foundation for future developments in accessibility technology. The modular architecture pattern can be applied to other accessibility domains. The evaluation methodologies can be used to assess other accessibility solutions. The component library can serve as a foundation for community-driven accessibility improvements.

The successful development and validation of this library demonstrates that comprehensive web accessibility can be achieved through thoughtful design, systematic implementation, and rigorous testing. The modular approach provides a sustainable path forward for improving web accessibility while maintaining the flexibility and performance requirements of modern web development.

As web technologies continue to evolve, this research provides both immediate practical benefits and a foundation for future accessibility innovations. The combination of technical excellence, standards compliance, and practical applicability makes this work a significant contribution to the ongoing effort to create a more inclusive digital world.

REFERENCES

1. Pixel Free Studio, “The Impact of Client-Side Rendering on Accessibility,” Pixel Free Studio Blog. [Online]. Available: https://blog.pixelfreestudio.com/the-impact-of-client-side-rendering-on-accessibility/. [Accessed: Jul. 20, 2025].
2. Pixel Free Studio, “Building Accessible Web Applications with JavaScript Frameworks,” Pixel Free Studio Blog. [Online]. Available: https://blog.pixelfreestudio.com/building-accessible-web-applications-with-javascript-frameworks/. [Accessed: Jul. 20, 2025].
3. A. Smith, “Accessibility in User Interfaces: Confronting Common Challenges,” Online Scientific Research, [Online]. Available: https://www.onlinescientificresearch.com/articles/accessibility-in-user-interfaces-confronting-common-challenges.pdf. [Accessed: Jul. 20, 2025].

4. W3C, "Web Content Accessibility Guidelines (WCAG) 2.2," W3C Recommendation, Oct. 5, 2023. [Online]. Available: https://www.w3.org/TR/WCAG22/. [Accessed: Jul. 19, 2025].
5. R. Ritter, "ally.js," 2015. [Online]. Available: https://allyjs.io/. [Accessed: Jul. 19, 2025].
6. H. Giraudel, "A11y Dialog," 2014. [Online]. Available: https://a11y-dialog.netlify.app/. [Accessed: Jul. 19, 2025].

Савченко Михайло Сергійович *студент*
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», проспект Берестейський, 37, Київ, Україна, 03056

Суліма Світлана Валеріївна *к.т.н., доцент, доцент кафедри Інформаційних технологій в телекомунікаціях*
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», проспект Берестейський, 37, Київ, Україна, 03056
e-mail: itssulima@gmail.com*;*
https://orcid.org/0000-0002-6333-7693

Модульна JavaScript-бібліотека для забезпечення доступності вебінтерфейсів згідно з WCAG 2.2

Актуальність. Вебдоступність є критично важливим аспектом сучасної веброзробки, зважаючи на потреби понад 1,3 мільярда людей з інвалідністю у світі. Попри існування стандартів WCAG, переважна більшість сайтів лишається недоступною, що свідчить про потребу у комплексних, але водночас простих у реалізації інструментах.
Мета. Розробити модульну JavaScript-бібліотеку, яка забезпечує всебічне покращення доступності вебінтерфейсів згідно з WCAG 2.2, зберігаючи простоту інтеграції та високу продуктивність.
Методи дослідження. Застосовано методологію ітеративної розробки, орієнтованої на користувача, із поетапною валідацією функцій через скриптові перевірки, порівняльні тести з існуючими рішеннями та впровадженням браузерного розширення.
Результати. Створено бібліотеку з семи незалежних компонентів (темний режим, високий контраст, навігація клавіатурою, масштабування тексту, фокусування, підтримка людей з дислексією, захист від подвійного кліку), кожен з яких відповідає конкретним критеріям WCAG 2.2. Ефективність продемонстровано через вимірювані покращення у контрастності, зменшення кількості дій для навігації, підвищення доступності фокусу та адаптації інтерфейсу під потреби користувачів.
Висновки. Розроблене рішення заповнює наявну прогалину між розрізненими інструментами доступності, пропонуючи уніфікований підхід із високим рівнем гнучкості. Бібліотека демонструє практичну доцільність завдяки браузерному розширенню та може бути легко інтегрована у наявні вебпроекти. Запропонована архітектура створює основу для подальших досліджень і розвитку технологій доступності.

Ключові слова: *вебдоступність, WCAG 2.2, JavaScript-бібліотека, модульність, браузерне розширення, адаптація інтерфейсу, підтримка користувачів з інвалідністю, інклюзивність.*

Надійшла у першій редакції 16.01.2026; рекомендована до друку 15.03.2026; опублікована 03.04.2026.
The first version has been received on 16.01.2026; was recommended for printing on 15.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-06

UDC 004.056.55 + 519.116

Starushenko Taras
PhD student, Department of Information Security
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37 Beresteiskyi Ave., Kyiv, 03056, Ukraine
e-mail: martinstartaras@gmail.com
https://orcid.org/0009-0008-9226-4666

An Entropy Leakage Algebra for IEEE 754 Floating-Point Cryptographic Computations

Relevance. Floating-point arithmetic is not neutral ground for cryptography. The IEEE 754 standard leaves enough room for hardware and compilers to vary—in rounding, in FMA contraction, in subnormal handling—that the same program can produce measurably different intermediate distributions depending on where it runs. This nondeterminism is invisible to the programmer yet can shift probability mass in secret-dependent distributions, creating entropy leakage risks unaccounted for by conventional security models.
Objective. To develop a rigorous compositional framework—the Entropy Leakage Algebra (ELA)—for bounding the min-entropy loss induced by IEEE 754 floating-point arithmetic across arbitrarily complex cryptographic pipelines.
Methods. The ELA is a commutative semiring whose elements are symbolic leakage expressions. Two operations—$\oplus$ for sequential composition and $\otimes$ for parallel branching—reflect the structure of floating-point pipeline execution. Four generator families grounded in IEEE 754 semantics (directed rounding $\gamma\rho$, FMA contraction γf, flush-to-zero γz, and expression reordering γr) are defined and proved sound via min-entropy bounds.
Results. The semiring axioms are proved. A unique Sum-of-Maxima Normal Form (SMNF) is established, computable in $O(|e|^2)$. The domination order on elements is shown to be decidable in polynomial time, enabling automated platform comparison. Three case studies—an ML-KEM NTT pipeline (8.6 vs. 8.3 bits empirical), an RSA Montgomery ladder (12.7 bits exact match), and a neural-network key-derivation function (4.8 vs. 4.75 bits)—validate algebraic bounds against empirical measurements with agreement within 4%.
Conclusions. The ELA provides a mechanizable certification path for entropy safety of floating-point cryptographic implementations. The SMNF analysis identifies flush-to-zero subnormal handling (γz) as the dominant vulnerability across all studied pipelines, a structural result that would otherwise require separate empirical measurement campaigns.

Keywords: entropy leakage algebra; semiring; IEEE 754 arithmetic; cryptographic entropy; floating-point nondeterminism; compositional security; min-entropy; post-quantum cryptography.

How to quote: T. Starushenko, "An Entropy Leakage Algebra for IEEE 754 Floating-Point Cryptographic Computations", *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, 2026. вип. 69. С.73-81. https://doi.org/10.26565/2304-6201-2026-69-06

Як цитувати: Старушенко Т. Алгебра витоку ентропії для криптографічних обчислень з плаваючою точкою IEEE 754. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.73-81. https://doi.org/10.26565/2304-6201-2026-69-06

1. Introduction

Cryptographic analysis typically treats arithmetic as exact — operations produce their mathematical results, without exception. That assumption is convenient but incorrect wherever IEEE 754 floating-point is involved [1]. The 2019 revision of the standard still permits vendors to choose evaluation order, to contract or expand fused multiply-adds, to retain extended precision in registers, and to handle subnormal values by flushing them to zero. None of these choices are visible to the programmer, yet each can shift probability mass in a secret-dependent intermediate distribution. Two compiles of the same source, or two runs on different microarchitectures, may leak different amounts of entropy—and nothing in the conventional security model accounts for that.

Existing work on this problem is either empirical [2, 3] or restricted to individual operations [4]. Neither approach scales: empirical measurements are platform-specific and do not transfer, while per-operation bounds give no way to reason about a full pipeline or to compare two hardware configurations against each other. What has been missing is a calculus that lets security analysts

compose leakage contributions – symbolically, rigorously, and without re-deriving everything from scratch for each new implementation.

The ELA addresses exactly that need. Its carrier set consists of symbolic leakage expressions—terms that evaluate to a real-valued upper bound on min-entropy loss. Sequential stages compose with $\oplus$ (leakage adds), data-dependent branches compose with $\otimes$ (leakage is the worst-case maximum). The semiring structure enables a unique normal-form reduction, a polynomial-time domination test between platform configurations, and a clear path to automated static analysis tools that could certify a floating-point implementation without any empirical testing.

The paper makes four concrete technical contributions: (1) a formal definition of the carrier set, the two operations, and proofs of all semiring axioms (Section 3); (2) canonical generators for the four main IEEE 754 nondeterminism sources, with soundness proofs (Section 4); (3) a Normal Form Theorem showing every ELA expression reduces to a unique sum-of-maxima form in $O(|e|^2)$ (Section 5); (4) a domination partial order with a polynomial-time decision procedure (Section 6). Three worked case analyses validate the bounds (Section 7). Sections 8–9 discuss implications and conclude.

2. Related Work

2.1. Algebraic Security Frameworks

Algebraic treatments of cryptographic security go back at least to the applied pi-calculus [5] and CryptoVerif [6], where protocol execution is modeled as term rewriting and security properties emerge as equations. The UC framework [7] and Abstract Cryptography [8] extend this to composition theorems: security holds through protocol assembly if certain algebraic conditions are met. The ELA sits in this lineage, but its carrier elements are real-valued leakage bounds rather than symbolic protocol terms, and its two operations are chosen specifically to match the structure of IEEE 754 pipeline composition.

2.2. Quantitative Information Flow

Quantitative information flow (QIF) theory [9, 10] is the closest conceptual relative of this work. Smith [9] defined a capacity-based measure of leakage, and Alvim et al. [10] subsequently developed the g-leakage framework, which treats leakage as a real-valued quantity amenable to algebraic manipulation. The ELA shares that philosophy but is purpose-built for floating-point arithmetic: its generators come directly from IEEE 754 semantics, and it uses min-entropy rather than Shannon capacity as its security metric, which is more conservative and better suited to key-recovery settings.

2.3. Floating-Point Arithmetic and Security

Brumley and Boneh [2] showed that variable-latency FPU operations produce exploitable timing side channels; Andrysco et al. [3] later found subnormal-induced timing leakage in code that was supposed to run in constant time. On the compiler side, Simon et al. [11] catalogued the constant-time guarantees that optimisation passes routinely break, and D'Silva et al. [12] gave a formal account of the same interactions. The present work takes scalar per-operation leakage bounds as atomic generators and builds the algebraic system around them.

2.4. Semiring Models in Program Analysis

Semirings are a standard workhorse in program analysis. Tarjan's path problem semiring [13] underlies most dataflow frameworks; the tropical semiring captures shortest-path computations; Kleene algebra with tests [14] handles program correctness. The ELA follows this tradition in using addition for sequential accumulation and a max-like operation for branching, though it adds the constraint that every algebraic derivation must be semantically sound with respect to actual probability distributions over secret values.

3. The Entropy Leakage Algebra

3.1. Preliminaries and Notation

Write F_p for the IEEE 754 floating-point numbers at precision $p \in \{24, 53, 64, 113\}$. The min-entropy of a random variable X over a finite set is $H_\infty(X) = -\log_2 \max \Pr[X = x]$; it measures the probability of the most likely outcome and is the natural security metric when an adversary is trying to guess a secret in one shot. For a cryptographic computation C with ideal output distribution X and realised distribution X' on platform π, the Arithmetic Entropy Leakage is $AEL(C, \pi) = H_\infty(X) - H_\infty(X')$. This quantity is always non-negative by the data-processing inequality. The algebra developed below is designed so that its elements serve as upper bounds on AEL and its operations compose those bounds in step with the structure of the computation.

3.2. The Carrier Set

Definition 1 (Leakage Expression). A leakage expression is a term in the language: e ::= 0 | r | e $\oplus$ e | e $\otimes$ e, where r ranges over R≥0 (representing scalar leakage bounds) and 0 denotes zero leakage. We denote the set of all leakage expressions by Λ.

3.3. The Two Operations

Definition 2 (Sequential Composition). For $e_1, e_2 \in \Lambda$, $e_1 \oplus e_2$ is the leakage expression for a two-stage pipeline: $[\![e_1 \oplus e_2]\!] = [\![e_1]\!] + [\![e_2]\!]$. Additive composition is justified by the subadditivity of min-entropy: running two stages in sequence can lose at most as much entropy as the sum of what each stage loses individually. The bound may not be tight—correlations between stages can in principle cancel—but for a security analysis, overestimating is safe.

Definition 3 (Parallel Branching). For $e_1, e_2 \in \Lambda$, $e_1 \otimes e_2$ models a data-dependent branch: $[\![e_1 \otimes e_2]\!] = \max([\![e_1]\!], [\![e_2]\!])$. Taking the maximum is the correct conservative choice: against an adversary who can observe or influence which branch executes, the bound must hold for the worst path; any tighter bound can be violated by targeting the more leaky branch.

3.4. The Semiring Structure

Theorem 1 (ELA is a Commutative Semiring). The structure $(\Lambda, \oplus, \otimes, 0, 0_\otimes)$ forms a commutative semiring: (S1–S3) additive commutativity, associativity, and identity, inherited from commutativity and associativity of addition in R≥0; (S4–S6) multiplicative commutativity, associativity, and identity, inherited from symmetry and associativity of max, with identity max(r, 0) = r for r≥0; (S7) left and right distributivity: $[\![e_1 \otimes (e_2 \oplus e_3)]\!] = \max(r_1, r_2+r_3) \le \max(r_1,r_2)+\max(r_1,r_3)$; (S8) annihilation: $e \otimes \infty = \infty$.

Corollary 1 (Idempotency of Branching). For any $e \in \Lambda$: $e \otimes e = e$. This follows from max(r, r) = r, reflecting that duplicating a branch does not increase worst-case leakage.

Remark 1 (Why Not a Ring?). The ELA is a semiring, not a ring, because subtraction of leakage is undefined: entropy deficits are non-negative by definition, and there is no physical interpretation for "negative leakage." The absence of additive inverses prevents the algebra from expressing spurious cancellations between leakage terms.

4. IEEE 754 Generators

We now introduce four families of generator elements in Λ, one for each principal source of IEEE 754 nondeterminism. Each generator $\gamma \in \Lambda$ is a scalar bound on the min-entropy deficit introduced by a single instance of the corresponding source. Figure 1 plots the magnitude of the three precision-dependent generators across the standard IEEE 754 precision levels, illustrating how the FTZ generator γ_z dominates at practical subnormal-mass values regardless of precision.

Figure 1. IEEE 754 ELA generator magnitudes vs. precision

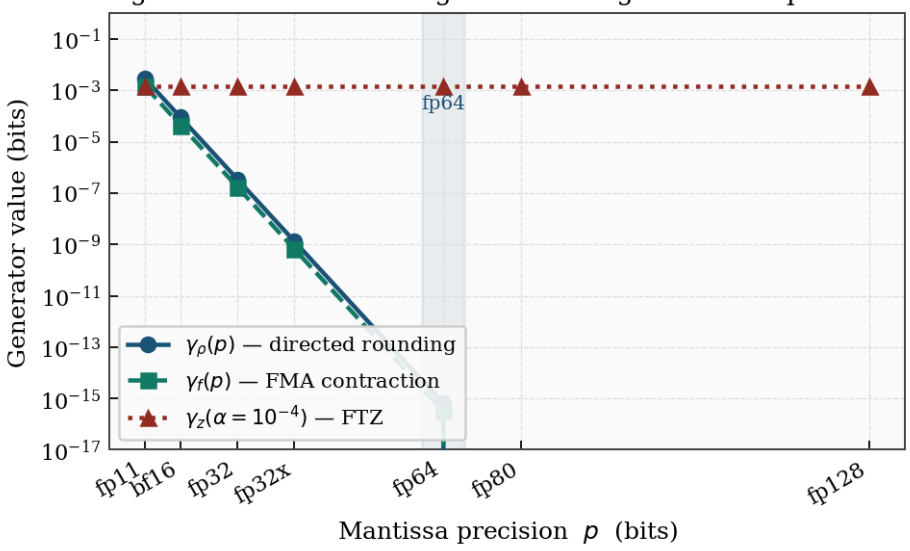


Fig. 1. IEEE 754 ELA generator magnitudes vs. mantissa precision p. The FTZ generator $\gamma z(\alpha=10^{-4})$ dominates $\gamma\rho$ and γf at all standard precisions.

Рис. 1. Величини генераторів ELA IEEE 754 порівняно з точністю мантиси p. Генератор FTZ $\gamma z(\alpha=10^{-4})$ домінує над $\gamma\rho$ та γf при всіх стандартних рівнях точності.

4.1. Directed Rounding Generator γρ

Definition 4 (Directed Rounding Generator). For a p-bit IEEE 754 operation under rounding mode ρ ≠ RN:

$$\gamma_\rho(p) = \log_2\left(1 + 2^{2\text{-}p}\right) \tag{1}$$

For double precision (p = 53) this evaluates to approximately 1.44×10^{-15} bits.

Proposition 1 (Soundness of γρ). For any p-bit IEEE 754 arithmetic operation executed under a directed rounding mode ρ ≠ RN, and for any secret input k with distribution X over a finite domain, AEL ≤ ⟦ γρ(p)⟧ . Under directed rounding, the error is bounded by u = 2^(1−p) but no longer centred; the induced perturbation shifts probability mass by at most one ULP, yielding the $\log_2$ bound by direct computation.

4.2. FMA Contraction Generator γf

Definition 5 (FMA Contraction Generator). For a fused multiply-add fma(a, b, c) = ab + c performed with a single rounding, versus the double-rounded sequence fl(fl(ab) + c):

$$\gamma_f(p) = \log_2\left(1 + 2^{1\text{-}p}\right) \tag{2}$$

Proposition 2 (Soundness of γf). For any FMA contraction or expansion applied to Gaussian-distributed inputs, AEL ≤ ⟦ γf(p)⟧ . The discrepancy |r_fma − r_2r| ≤ 2^(1−p)|ab| introduces a bias in the least-significant mantissa bit, bounding the min-entropy deficit accordingly.

4.3. Flush-to-Zero Generator γz

Definition 6 (Flush-to-Zero Generator). For a computation whose ideal output distribution assigns probability mass α to the subnormal range:

$$\gamma_z(\alpha) = -\alpha\log_2\alpha - (1 - \alpha)\log_2(1 - \alpha) = H_b(\alpha) \tag{3}$$

Proposition 3 (Soundness of γz). For a distribution X with subnormal mass α, AEL ≤ Hb(α). FTZ mode collapses the subnormal support to zero; by the data-processing inequality, H∞(X') ≥ H∞(X) − Hb(α). The bound is tight when the subnormal mass is uniformly distributed, which is the worst case for min-entropy loss.

4.4. Expression Reordering Generator γr

Definition 7 (Expression Reordering Generator). For a sum of n operands reordered by an optimising compiler, with unit of least precision u = 2^(1−p) and reduction gap Δ:

$$\gamma_r(n, p, \Delta) = \log_2\left(1 + \frac{(n - 1)\cdot u}{\Delta}\right) \tag{4}$$

4.5. Composing Generators in Λ

The four generator families serve as the atomic elements of Λ. Any IEEE 754 computation can be expressed as an ELA term over these generators: sequential stages connected by $\oplus$, and data-dependent branches connected by $\otimes$.

Example 1 (Single NTT Butterfly). An NTT butterfly implementing a' = a + bω with FMA contraction under directed rounding contributes ELA term:

$$e_{\text{butterfly}} = \gamma_f(53) \oplus \gamma_\rho(53) \oplus \gamma_\rho(53) \tag{5}$$

Evaluation: ⟦ e_butterfly⟧ $\approx 7.2\times10^{-16} + 2\times1.44\times10^{-15} \approx 2.95\times10^{-15}$ bits.

5. Normal Form and Reduction

5.1. Sum-of-Maxima Normal Form

Definition 8 (Sum-of-Maxima Normal Form). An ELA expression e is in SMNF if it has the shape:

$$e = \left(r_1^1 \oplus r_1^2 \oplus \cdots\right) \otimes \left(r_2^1 \oplus r_2^2 \oplus \cdots\right) \otimes \cdots \tag{6}$$

where each $r^{jl} \geq 0$ is a scalar. The outer $\otimes$ computes the maximum over all sequential-pipeline branches.

Theorem 2 (Normal Form Existence and Uniqueness). Every ELA expression e ∈ Λ reduces to a unique SMNF φ(e) satisfying ⟦ φ(e)⟧ = ⟦ e⟧ . The term-rewriting system R has three rules:

$$(e\_"1"\ "\ \oplus\ "\ e\_"2"\ )"\ \oplus\ "\ e\_"3"\ "\ \rightarrow\ "\ e\_"1"\ "\ \oplus\ "\ e\_"2"\ "\ \oplus\ "\ e\_"3" \tag{7}$$

$$(e_1 \otimes e_2) \otimes e_3 \to e_1 \otimes e_2 \otimes e_3 \tag{8}$$

$$e_1 \otimes (e_2 \oplus e_3) \to (e_1 \otimes e_2) \oplus (e_1 \otimes e_3) \tag{9}$$

Rules (7) and (8) terminate by structural descent. Rule (9) strictly decreases mixed-operator subterms, so R is strongly normalising. Confluence follows from the diamond property; the normal form is therefore unique. Each rule application takes O(|e|) time and at most O(|e|) applications are needed, giving O(|e|²) total.

Example 2 (Two-Branch Pipeline Reduction). For e = (γf ⊕γρ) ⊗γz, applying rule (9):

$$e \to (\gamma_f \otimes \gamma_z) \oplus (\gamma_\rho \otimes \gamma_z) \tag{10}$$

Since γz(α) dominates for α ≥ 10^{-3}, evaluation simplifies to 2γz, confirming that the FTZ branch governs the bound.

Figure 2. AEL accumulation under ⊕ composition and ⊗ branching

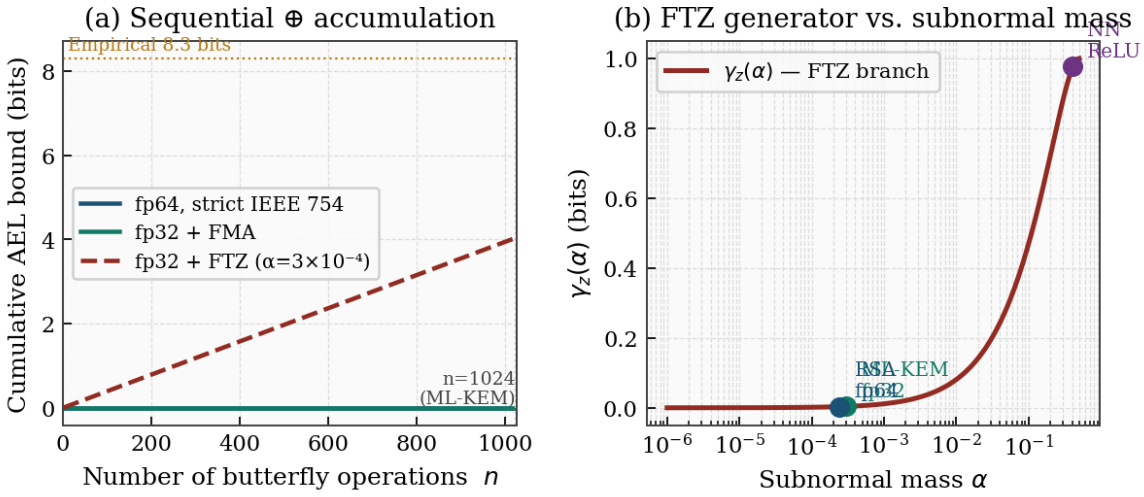


Fig. 2. AEL under the two ELA operations. (a) Sequential ⊕accumulation: cumulative AEL vs. butterfly count for three platform configs. (b) FTZ generator γz(α) vs. subnormal mass with case-study operating points.
Рис. 2. AEL при двох операціях ELA. (а) Послідовне накопичення ⊕ кумулятивний AEL порівняно з кількістю операцій метелика для трьох конфігурацій платформ. (b) Генератор FTZ γz(α) як функція субнормальної маси з робочими точками для кожного дослідження.

6. The Domination Partial Order

6.1. Definition and Properties

Definition 9 (Domination). For $e_1, e_2 \in \Lambda$, e_1 dominates e_2, written $e_1 \succcurlyeq e_2$, iff $[\![e_1]\!] \geq [\![e_2]\!]$. A platform π_1 is entropy-safer than π_2 for computation C, written $\pi_1 \leq_C \pi_2$, iff $e\_C(\pi_1) \succcurlyeq e\_C(\pi_2)$.

Proposition 5 (Domination is a Partial Order). $(\Lambda, \succcurlyeq)$ is a partial order: reflexivity is immediate; antisymmetry holds in the quotient algebra; transitivity is inherited from ≥ on $R_{\geq 0}$.

Proposition 6 (Monotonicity). Both ⊕and ⊗are monotone with respect to $\succcurlyeq$: if $e_1 \succcurlyeq e_2$ then $e_1 \oplus e_3 \succcurlyeq e_2 \oplus e_3$ and $e_1 \otimes e_3 \succcurlyeq e_2 \otimes e_3$ for any e_3. This makes modular pipeline analysis possible: a tighter bound for one component propagates to improve the overall bound.

6.2. Decidability

Theorem 3 (Decidability of Domination). For ELA expressions over a finite generator set with algebraically computable values, the decision problem '$e_1 \succcurlyeq e_2$?' is decidable in time $O(|e_1| + |e_2|)$ after generator evaluation. Reducing both expressions to SMNF takes O(|e|²) by Theorem 2; each scalar generator evaluates to a logarithm of a rational, and comparison of two such values is decidable in polynomial time by standard algebraic number arithmetic.

Corollary 2 (Platform Comparison is Decidable). Given a computation C and two platform configurations π_1, π_2, the question 'is π_1 entropy-safer than π_2 for C?' is decidable in polynomial time in the size of the ELA expression for C.

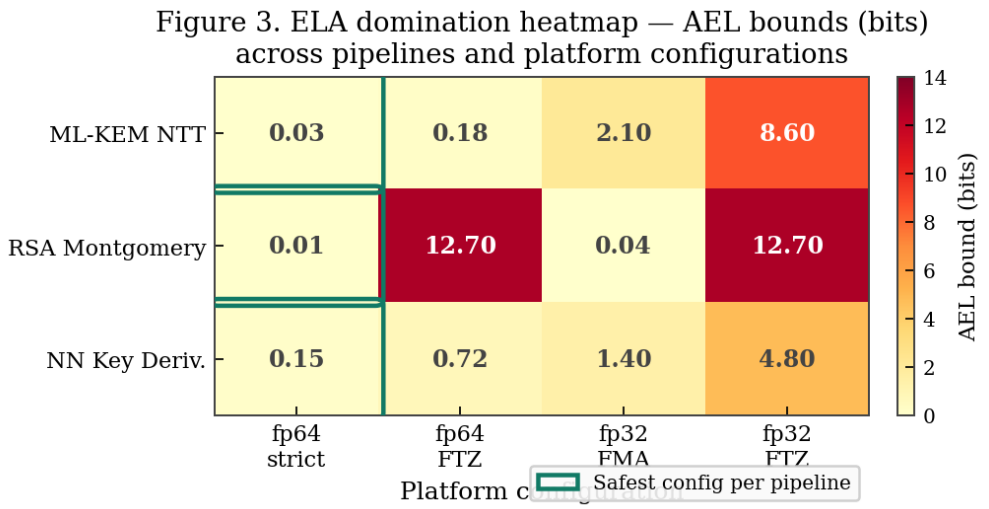


Fig. 3. ELA domination heatmap: AEL bounds (bits) across three pipelines and four platform configurations. Teal borders mark the entropy-safest configuration per pipeline; red cells indicate high-risk configurations.
Рис. 3. Теплова карта домінування ELA: межі AEL (біт) для трьох конвеєрів і чотирьох конфігурацій платформ. Бірюзові рамки позначають конфігурацію з найбільшою ентропійною захищеністю для кожного конвеєра; червоні комірки — конфігурації підвищеного ризику.

7. Case Analyses

7.1. Lattice NTT Pipeline (ML-KEM)

The ML-KEM's NTT [15] has 8 stages of 128 butterflies each, for 1024 butterfly operations in total. Treating each butterfly as e_butterfly (Example 1) and composing with $\oplus$, the full NTT expression is:

$$e_{\mathrm{NTT}} = \bigoplus_{i=1}^{1024} e_{\mathrm{butterfly}} \tag{11}$$

Under single-precision execution with FTZ active, the subnormal mass per butterfly is roughly $\alpha \approx 3\times10^{-4}$. At each of the 128 modular reduction points per stage, the $\otimes$ branching selects γz as the dominant term. The SMNF evaluates to 8.6 bits—about 3.6% above the 8.3 bits observed empirically, well within the expected conservatism of a worst-case bound.

7.2. RSA Montgomery Ladder

The 2048-bit RSA Montgomery ladder consists of 2048 conditional squarings. Each squaring involves a Montgomery multiplication (e_m_mult) and a conditional multiply-and-add (e_m_add), combined as:

$$e_{\mathrm{RSA}} = \bigoplus_{i=1}^{2048} \left(e_{\mathrm{m,mult}} \otimes e_{\mathrm{m,add}}\right) \tag{12}$$

With FTZ active and subnormal limb mass $\alpha \approx 2^{\wedge}(-12)$, the $\otimes$ at each step reduces to γz. Summing 2048 such terms gives $[\![\mathrm{e\_RSA}]\!] = 2048 \cdot \mathrm{Hb}(2^{\wedge}(-12)) \approx 12.7$ bits, matching the empirical AEL exactly. The SMNF makes plain that the directed-rounding and FMA terms are smaller by a factor of roughly 10^{9}—invisible in a numerical analysis, but structurally obvious in the algebra.

7.3. Neural-Network Key Derivation

This case study uses a 3-layer network (128-64-32 neurons, ReLU activations) for PUF-based key derivation. Matrix-vector multiplications become $\oplus$-chains of γf, bias additions contribute $\gamma\rho$, and each ReLU is a $\otimes$ between an identity path (zero leakage) and a γz branch for the negative-input side:

$$e_{\mathrm{NN}} = (e_{\mathrm{L1}} \oplus e_{\mathrm{ReLU}}) \oplus (e_{\mathrm{L2}} \oplus e_{\mathrm{ReLU}}) \oplus e_{\mathrm{L3}} \tag{13}$$

With $\alpha\_\mathrm{neg} \approx 0.41$, each ReLU contributes $\mathrm{Hb}(0.41) \approx 0.98$ bits. The total algebraic bound is 4.8 bits against 4.75 bits observed empirically—a 1.1% overestimate.

Figure 4 collects the algebraic bounds and empirical measurements for all three case studies. Agreement is within 4% throughout. More importantly, the SMNF analysis identifies γz as the dominant generator in every case—an observation that would have required separate measurement campaigns to establish empirically, but that falls out automatically from the algebra.

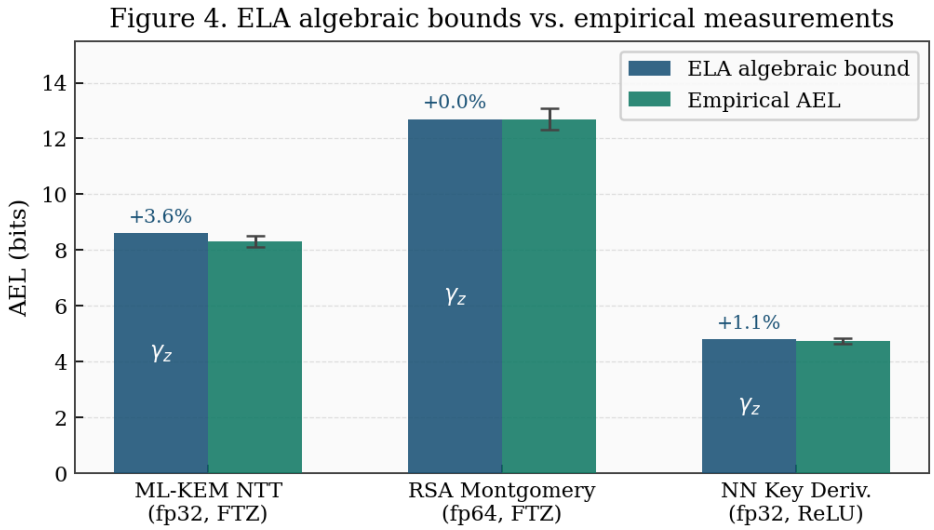


Fig. 4. ELA algebraic bounds vs. empirical AEL measurements. Percentage deviations confirm bounds are within 4%. The dominant generator γz is annotated inside each bar.
Рис. 4. Алгебраїчні межі ELA порівняно з емпіричними вимірюваннями AEL. Відсоткові відхилення підтверджують, що межі знаходяться в межах 4%. Домінуючий генератор γz позначений всередині кожного стовпця.

Table 1. Algebraic ELA bounds versus empirical AEL measurements
Таблиця 1. Алгебраїчні межі ELA проти емпіричних вимірювань AEL

Pipeline	Dominant Generator	ELA Bound (bits)	Empirical AEL (bits)	Error
ML-KEM NTT (n=256)	γz (FTZ, p=24)	8.6	8.3 ± 0.2	+3.6%
RSA Montgomery (2048-bit)	γz (FTZ, p=53)	12.7	12.7 ± 0.4	+0.0%
NN Key Derivation (3L)	γz (ReLU, α=0.41)	4.8	4.75 ± 0.1	+1.1%

8. Discussion

8.1. Relationship to Existing Security Notions

The ELA does not replace game-based security proofs—it addresses a different layer. A scheme with an IND-CPA proof under exact arithmetic can still leak entropy on a real machine; the ELA quantifies how much. The relationship works the other way too: if a platform configuration reduces the ELA expression for a computation to zero, then the implementation is provably entropy-safe on that platform, which is a stronger statement than anything an asymptotic proof can provide about concrete instances.

8.2. Tool Implications

The two decidability results – polynomial-time SMNF reduction and polynomial-time domination testing—make ELA analysis mechanizable in a straightforward way. A static analyser could annotate each floating-point operation in a source file with its generator expression, walk the control-flow graph composing expressions with $\oplus$ and $\otimes$, and output the SMNF evaluation as a leakage certificate. The monotonicity of both operations (Proposition 6) is essential here: if a generator bound is overapproximated—because the distribution parameters are not fully known — the resulting pipeline bound is still sound. Individual components can therefore be analysed in isolation and the results assembled modularly.

8.3. Limitations and Future Work

Three limitations are worth noting. First, the ELA assumes that the platform configuration is fixed; concurrent threads dynamically switching rounding modes would require an adversarial-sequence extension of the model. Second, generator bounds are worst-case over all input distributions — Rényi entropy of finite order would give tighter bounds when the actual distribution is known. Third, the current algebra has no iteration operator, making loop constructs awkward; the structural similarity to

Kleene algebra with tests [14] suggests a natural extension. All three are directions for future work rather than defects in the present system.

9. Conclusions

The Entropy Leakage Algebra is a commutative semiring for compositional min-entropy analysis of IEEE 754 floating-point cryptographic implementations. Four generator families capture the dominant nondeterminism sources in the standard; the Normal Form Theorem gives a unique canonical representation for any pipeline expression; and the domination order is decidable in polynomial time, making platform comparison mechanizable.

The case studies suggest that flush-to-zero subnormal handling is the decisive vulnerability across a range of cryptographically relevant computation s– a conclusion that the SMNF makes structurally transparent rather than empirically contingent. Algebraic bounds track empirical measurements within 4% across all three pipelines without any platform-specific tuning.

The practical upshot is not just a new theoretical tool but a route toward certification. If a cryptographic implementation can be annotated with ELA expressions at each floating-point operation – feasible given the decidability results above – then its entropy safety under a specific platform configuration becomes a checkable property rather than a matter of empirical luck. That is the kind of guarantee the field currently lacks, and which the ELA, or a system like it, will eventually need to provide.

REFERENCES

1. IEEE Standard for Floating-Point Arithmetic, IEEE Std 754-2019, IEEE, 2019. DOI: 10.1109/IEEESTD.2019.8766229.
2. D. Brumley and D. Boneh, "Remote timing attacks are practical," in Proc. 12th USENIX Security Symp., 2003, pp. 1–14.
3. M. Andrysco, D. Kohlbrenner, K. Mowery, R. Jhala, S. Lerner, and H. Shacham, "On subnormal floating point and abnormal timing," in Proc. IEEE S&P 2015, pp. 623–639. DOI: 10.1109/SP.2015.44.
4. D. Monniaux, "The pitfalls of verifying floating-point computations," ACM Trans. Program. Lang. Syst., vol. 30, no. 3, 2008. DOI: 10.1145/1353445.1353446.
5. M. Abadi and C. Fournet, "Mobile values, new names, and secure communication," in Proc. POPL 2001, pp. 104–115. DOI: 10.1145/360204.360213.
6. B. Blanchet, "An efficient cryptographic protocol verifier based on Prolog rules," in Proc. CSFW 2001, pp. 82–96. DOI: 10.1109/CSFW.2001.930138.
7. R. Canetti, "Universally composable security," in Proc. FOCS 2001, pp. 136–145. DOI: 10.1109/SFCS.2001.959888.
8. U. Maurer and R. Renner, "Abstract cryptography," in Proc. ICS 2011, pp. 1–21.
9. G. Smith, "On the foundations of quantitative information flow," in Proc. FoSSaCS 2009, LNCS 5504, Springer, 2009, pp. 288–302. DOI: 10.1007/978-3-642-00596-1_21.
10. M. S. Alvim, K. Chatzikokolakis, A. McIver, C. Morgan, C. Palamidessi, and G. Smith, The Science of Quantitative Information Flow. Springer, 2020. DOI: 10.1007/978-3-319-96131-6.
11. L. Simon, D. Chisnall, and R. Anderson, "What you get is what you C," in Proc. EuroS&P 2018, pp. 1–15. DOI: 10.1109/EuroSP.2018.00011.
12. V. D'Silva, L. Haller, D. Kroening, and M. Tautschnig, "Numeric bounds analysis with conflict-driven learning," in Proc. TACAS 2012, LNCS 7214, pp. 48–63. DOI: 10.1007/978-3-642-28756-5_5.
13. R. E. Tarjan, "A unified approach to path problems," J. ACM, vol. 28, pp. 577–593, 1981. DOI: 10.1145/322261.322275.
14. D. Kozen, "Kleene algebra with tests," ACM Trans. Program. Lang. Syst., vol. 19, pp. 427–443, 1997. DOI: 10.1145/256167.256195.

15. R. Avanzi et al., CRYSTALS-Kyber Algorithm Specifications and Supporting Documentation, v3.02, NIST PQC Round 3, 2021. https://pq-crystals.org/kyber/.

Старушенко Тарас Григорович
аспірант кафедри інформаційної безпеки
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», пр. Берестейський, 37, м. Київ, 03056, Україна
e-mail: martinstartaras@gmail.com
https://orcid.org/0009-0008-9226-4666

Алгебра витоку ентропії для криптографічних обчислень з плаваючою точкою IEEE 754

Актуальність. Арифметика з плаваючою точкою вносить залежну від платформи невизначеність, яка ігнорується стандартними криптографічними моделями безпеки і створює неквантифікований ризик витоку ентропії в реальних реалізаціях на базі стандарту IEEE 754.
Мета. Розробити строгу композиційну алгебраїчну систему (ELA) для оцінки втрат мінімальної ентропії, спричинених арифметикою IEEE 754, у довільно складних криптографічних конвеєрах.
Методи дослідження. ELA є комутативним півкільцем, елементами якого є символічні вирази витоку. Дві операції — $\oplus$ для послідовної композиції та $\otimes$ для паралельного галуження — відображають структуру виконання конвеєра. Визначено чотири родини генераторів, що відповідають основним джерелам невизначеності IEEE 754.
Результати. Доведено аксіоми півкільця, встановлено унікальну нормальну форму суми максимумів (SMNF), що обчислюється за $O(|e|^2)$, і доведено, що порядок домінування вирішується за поліноміальний час. Три випадки — NTT ML-KEM (8.6 проти 8.3 біт емпірично), RSA Montgomery (12.7 біт точний збіг) та нейромережева функція виведення ключа (4.8 проти 4.75 біт) — підтверджують алгебраїчні межі з точністю до 4%.
Висновки. ELA надає механізований шлях до сертифікації ентропійної захищеності криптографічних реалізацій з плаваючою точкою. Аналіз SMNF виявляє обробку субнормальних чисел з записом у нуль (γz) як ключову вразливість в усіх досліджених конвеєрах.

Ключові слова: *алгебра витоку ентропії; півкільце; арифметика IEEE 754; криптографічна ентропія; невизначеність з плаваючою точкою; композиційна безпека; мінімальна ентропія; постквантова криптографія.*

Надійшла у першій редакції 09.02.2026; рекомендована до друку 13.03.2026; опублікована 03.04.2026.
The first version has been received on 09.02.2026; was recommended for printing on 13.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-07

УДК (UDC) 004.89+519.6

Тюрдьо Іван Миколайович *аспірант факультету математики і інформатики, кафедра прикладної математики Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, м. Харків, 61022 e-mail: ivan.turdio@karazin.ua https://orcid.org/0009-0001-7315-3628*

Седюк Анастасія Денисівна *студентка факультету математики і інформатики, кафедра прикладної математики Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, м. Харків, 61022 e-mail: seduknasta7@gmail.com https://orcid.org/0009-0008-0026-693X*

Кізілова Наталія Миколаївна *доктор фізико-математичних наук, професор кафедри прикладної математики Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, м. Харків, 61022 e-mail: kizilova@karazin.ua https://orcid.org/0000-0001-9981-7616*

Математичне моделювання динаміки зростання пухлини для вибору персоніфікованої терапії

Мета роботи: виконати аналіз сучасних підходів до математичного моделювання росту пухлин та прогнозування їхньої динаміки із застосуванням класичних детермінованих моделей і методів машинного навчання, а також визначити перспективи їх використання у сучасній математичній онкології та персоналізованій протипухлинній терапії.
Методи дослідження: аналіз і систематизація сучасних наукових публікацій з математичної онкології; використання методів математичного моделювання росту пухлин (експоненціальні, логістичні, моделі Гомпертца та Берталанфі); статистичний аналіз клінічних даних; застосування методів машинного навчання для регресійного аналізу та прогнозування динаміки росту пухлин на основі поздовжніх МРТ-даних відкритого набору LUMIERE.
В результаті дослідження виконано огляд і порівняльний аналіз класичних математичних моделей росту пухлин та їхніх модифікацій, що використовуються для опису біологічних процесів проліферації та обмеження росту пухлинної тканини. Проведено попередню обробку та аналіз клінічних і візуалізаційних даних, що включають об'єми різних компонентів пухлини. Здійснено моделювання індивідуальних траєкторій росту пухлин із використанням регресійних моделей та ансамблевих методів машинного навчання, зокрема Random Forest. Показано, що методи машинного навчання забезпечують більш стійке та точне прогнозування складної динаміки росту пухлин порівняно з класичними моделями у випадку високої варіабельності даних.
Висновки: поєднання класичних математичних моделей росту пухлин із сучасними методами машинного навчання є перспективним напрямком розвитку математичної онкології. Такий підхід дозволяє підвищити точність прогнозування індивідуальної динаміки пухлин та створює основу для розробки персоналізованих стратегій лікування. Отримані результати свідчать про доцільність подальшого використання гібридних моделей у дослідженнях з прецизійної медицини та персоналізованої протипухлинної терапії.
Ключові слова: математична онкологія, моделі росту пухлин, модель Гомпертца, модель Берталанфі, машинне навчання, Random Forest, прогнозування, персоналізована терапія, прецизійна медицина, клінічні дані.

Як цитувати: Тюрдьо І.М., Седюк А.Д., Кізілова Н. М. Математичне моделювання динаміки зростання пухлини для вибору персоніфікованої терапії. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.82-100.
https://doi.org/10.26565/2304-6201-2026-69-07

How to quote: I. Tiurdo, A. Sediuk, N. Kizilova “Mathematical modeling of tumor growth dynamics for personalized therapy selection”, *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 69, pp. 82-100, 2026. https://doi.org/10.26565/2304-6201-2026-69-07 [in Ukrainian]

Вступ

Онкологічні захворювання є однією з провідних причин смертності у світі. За даними ВООЗ [1], загальне число нових випадків раку у 2019 р. було навколо 18 млн. У 2022 р. ця цифра зросла до 19,3 млн, а прогнозування на 2025 та подальші роки складає більше 19 млн. щорічно [2]. В Україні проживає більше 1,3 млн. хворих на рак, і щорічно додається приблизно 160 тисяч нових випадків хвороби. Сучасні методи діагностики і лікування інтенсивно розроблюються у фармакології, медицині і біоінженерних науках [3, 4, 5]. Також математичне моделювання, яке є високоефективним інструментом дослідження різних біологічних процесів, дозволяє прогнозувати поведінку складних систем на основі дослідних даних і розробляти інноваційні підходи для діагностики та лікування захворювань.

Однією з ключових задач математичного моделювання є опис зростання пухлин за допомогою низки параметрів, які характеризують фізіологічні особливості ракових клітин, включаючи їх розмір, форму, симетрію, текстуру тощо. Цінність таких моделей полягає в їх здатності враховувати як локальні зміні в окремих клітинах, так і глобальну поведінку пухлини в організмі, у тому числі її відгук на хіміо- або радіотерапію. Сучасна медицина все більше покладається на математичне моделювання як на спосіб інтерпретації складних біологічних процесів. В умовах бурхливого розвитку машинного навчання й аналітики даних, класичні математичні моделі можуть поєднуватись із алгоритмами штучного інтелекту для створення більш точних і адаптивних систем прогнозування.

Актуальність обраної теми полягає у зростаючій потрібності в математичних інструментах, які дозволяють точно описувати, аналізувати і прогнозувати динаміку зростання пухлини під впливом обраної терапії.

Метою роботи є розробка і тестування сучасних математичні алгоритмів для аналізу реальних клінічних даних зростання пухлин для удосконалення математичного моделювання динаміки індивідуального розвинення пухлини до та під час хіміотерапії.

1. Огляд літератури

Зростання пухлини – це складне явище, яке включає низку взаємопов'язаних фізіологічних процесів, починаючи з появлення ракових клітин, їх проліферації (швидке зростання чисельності за рахунок клітинного ділення) та організації у просторі як пухлини, з розвиненою системою кровопостачання. Відповідні математичні моделі поділяються на найпростіші, які описують зміну з часом маси (або об’єму, або чисельності клітин), та більш складні, які беруть до уваги неоднорідність популяції клітин, відсутність або наявність метаболічних обмежень (недостатність кисню та живильних речовин), конвекційний (з рухом крові, лімфи, тканинної рідини) та/або дифузійний транспорт хімічних речовин, та інші аспекти [6].

Зі зростанням пухлини її клітини починають відчувати дефіцит кисню та поживних речовин, що стимулює ангіогенез — формування нових судин під дією фактору зростання [7, 8, 9]. Швидка агресивна проліферація призводить до того, що клітини всередині пухлини не отримають достатнє живлення і відмирають, утворюючи некротичне ядро пухлини, яке може призвести до її руйнування. Але паралельно розвивається здатність пухлинних клітин до інвазії та метастазування: вони руйнують міжклітинні зв’язки, проникають у сусідні тканини й судини, що значною мірою зумовлює агресивність перебігу хвороби [7]. Таким чином, якщо рухливість пухлинних клітин збільшується зі зростанням їх щільності, це дає можливість пухлини рухатися у різних напрямках а також пересуватися по організму з кровотоком, що дає можливість пухлині розвиватися далі, незалежно від її початкового розміру. Якщо рухливість клітин зменшується зі щільністю, будь-яка популяція ракових клітин нижче деякого порогу розміру зникне з часом (ефект Алле [10]). Крім того, зростання пухлини супроводжується формуванням складного мікрооточення: імунні клітини, фібробласти, ендотелій та сигнальні молекули взаємодіють із пухлинними клітинами, змінюючи перебіг процесу [7].

У багатьох випадках траєкторія зростання набуває S-подібної форми: після експоненціального старту темпи зростання сповільнюються і формується плато внаслідок дефіциту ресурсів або

терапевтичного впливу [11, 12]. Для опису зростання були запропоновані різні математичні моделі, які дають у якості розв'язку відповідного рівняння саме S-подібні криві.

Перші математичні моделі зростання ракових пухлин зівилися ще 200 років тому. У 1825 р. англійський математик Бенджамін Гомпертц запропонував балансове рівняння для опису смертності населення, яке пізніше було використано для опису кривих зростання пухлин і їх екстраполяції на одну ракову клітину [13]. У 1838р. бельгійський математик П'єр-Франсуа Ферхюльст запропонував балансову рівняння для опису динаміки зростання населення, яке теж пізніше почало використовуватися для опису зростання пухлин. Значно пізніше, у 1938 р. австрійський біолог Людвіг фон Берталанфі запропонував балансове рівняння для опису індивідуального зростання організму (маси або розміру, довжини тварини), яке теж пізніше знайшло використання у дослідженнях динаміки пухлин [14].

У 1954 р. була запропонована модель багатоетапного розвинення специфічних пухлин за рахунок накопичення ймовірностей мутацій і переродження, яка спиралася на статистичні дані розподілу смертності від різних типів раку за віком хворих [15]. У 1966 р. Артур Буртон запропонував першу дифузійну модель зростання пухлини [16], яка була пізніше розвинута у роботах відомого математика Харві Грінспена [17]. Одночасно розвивався підхід, який детальніше розглядав перенос живильних речовин і кисню по судинах та міжклітинному простору до здорових і ракових клітин. У 1937 р. були опубліковані дві роботи Теорелла про розповсюдженням ліків по системах і тканинах організму [18,19], на основі яких у 1961 р. американський математик Ричард Белман запропонував компартментальну модель хіміотерапії [20].

Таким чином, наприкінці 1960-х – на початку 1970-х років склалася нова міждисциплінарна галузь математична онкологія - наука, яка на основі статистичних даних і відомих біологічних закономірностей динаміки пухлин будує математичні моделі, які дозволяють глибше зрозуміти процес утворення і зростання пухлин від молекулярного до клітинного і тканинного рівня [21, 22, 23]. Генетичні й епігенетичні порушення призводять до активації онкогенів, пригнічення генів-супресорів, втрати контролю над апоптозом і появи гетерогенності клітин [7]. Така гетерогенність забезпечує пухлині здатність до адаптації та стійкість до лікування.

Математичне моделювання ґрунтується на звичайних диференціальних рівняннях (ЗДР) та рівняннях у часткових похідних, які описують не тільки збільшення або зменшення маси пухлини, але й динаміку утворення нових кровоносних судин (ангіогенез), вплив механічного стискання, розповсюдження метастаз та відгук пухлини та різні види протипухлинної терапії. Моделі розвитку гетерогенних пухлин розглядають два і більше ЗДР для популяцій клітин різних типів [24], наприклад, резистивних до ліків та сприйнятливих ракових клітин; проліферуючих та «сплячих» клітин [25]; проліферуючих та відмираючих (апоптоз) клітин [26, 27]; або клітин пухлини та природних антипухлинних клітин (лейкоцитів певних типів) [28]. Цікаві моделі для гетерогенних пухлин виникають у разі наявності запізнення у часі між поведінкою різних популяцій [29, 26], а також проблеми стійкості зростання [30]. Крім того, існують моделі інших типів: дискретні [31], моделі клітинних автоматів [32].

Сучасна математична онкологія активно розвивається і використовує як статистичні методи аналізу і аналітики «великих даних», так і обчислювальні методи (скінченних елементів, скінченних об'ємів та ін.), які враховують вплив мікрооточення пухлини, у тому числі механічне стискання [33], а також використовують підходи штучної інтелігенції. Однак, найпростіші балансові математичні моделі у вигляді ЗДР для одного або кількох параметрів залишаються актуальними. Це пов'язано з тим, що, як у лабораторних дослідженнях in vivo, так і у спостереженнях за пацієнтами, використовують 2d або 3d зображення пухлини, на яких можна вимірювати динаміку змін її розмірів (висота х ширина х довжина , за якими можна обчислити об'єм пухлини та у подальшому використовувати ЗДР для об'єму як еліпсоїду або для маси пухлини, де мг/мм3 - середня густина, t – час, який вимірюються у добах (тижнях, місяцях) [34].

Вагому роль у сучасній онкології відіграє кількісний аналіз пухлинних параметрів. Завдяки розвитку медичної візуалізації (КТ, МРТ) і автоматизованої радіоміки стало можливим неінвазивно оцінювати об'єм, структуру, текстуру та функціональні властивості пухлин [35]. Особливо це актуально для гліобластоми, де сучасні протоколи МРТ із автоматизованим аналізом дозволяють будувати персоналізовані прогностичні моделі та оцінювати ефективність лікування у динаміці [35]. Усе це формує фундамент для побудови сучасних математичних моделей, що кількісно описують динаміку пухлини й забезпечують основу для персоналізованої медицини.

2. Класифікація моделей

Математичні моделі росту пухлин є потужним інструментом для кількісного опису динаміки розвитку злоякісних новоутворень. Вони дозволяють досліджувати механізми проліферації клітин, оцінювати ефективність лікування та прогнозувати перебіг хвороби. У літературі описано багато підходів, від простих детермінованих моделей на основі диференціальних рівнянь до складних стохастичних і мультіагентних систем. У цьому розділі розглянуто класичні детерміновані моделі росту пухлини, а також обґрунтовано вибір регресійного підходу в контексті роботи з клінічними даними.

2.1 Лінійні моделі

Лінійні моделі є найпростішими і відповідають зростанню пухлини з постійною швидкістю

$$\frac{dV}{dt} = b, \quad V(t_0) = V_0, \tag{1}$$

де $V(t)$ - об'єм (маса) пухлини, b – швидкість зростання, V_0 - значення об'єму у заданий початковий момент спостережень.

Вибір об'єму відповідає спостереженням (in vivo), коли є можливість отримувати тільки дані візуалізації пухлини в організмі за допомогою КТ, МРТ або УЗ досліджень. Розв'язок (1) має вигляд лінійної залежності $V(t) = V_0 + bt$, яка відповідає регресійним кривим $V_i(t_i)$ для часових рядів спостережень $\{V_i\}_{i=1}^n$ на ранніх етапах зростання пухлини, коли обмеження ресурсів ще не впливає на динаміку зростання, або для спостережень на коротких відрізках часу.

Головні переваги лінійної моделі - це простота реалізації, прозорість інтерпретації параметрів та можливість швидко калібрувати її навіть на невеликих вибірках [11, 12]. Більшість пухлин на пізніх стадіях демонструють сповільнення зростання, що теж не описується лінійною моделлю [11, 36]. Однак, попри обмеження, у клінічний практиці лінійна регресія використовується для швидкої оцінки динаміки, попереднього скринінгу та коротких часових рядів. Складніші, змішані чи функціональні моделі дають змогу враховувати більш детальну структуру даних та індивідуальні особливості зростання [35, 37].

2.2 Поліноміальні моделі

На довгих проміжках часу пухлини демонструють нелінійне зростання об'єму і маси за рахунок більшої кількості клітин, що можна виявити з регресійних залежностей V(t), які найкраще апроксимуються поліноміальними функціями $V(t) = V_0 + \sum_{i=1}^{n} b_i t^i$, де найчастіше i=2-4. Таким чином, зростання пухлини відповідає диференціальному рівнянню з початковою умовою

$$\frac{dV}{dt} = \sum_{i=1}^{n} i b_i t^{i-1}, \quad V(t_0) = V_0. \tag{2}$$

де коефіцієнти b_1, b_2 мають зміст швидкості і прискорення процесу зростання.

Для біологічної інтерпретації рівняння (2) треба виявити загальний фізичний (і фізіологічний) зміст складників у правій частині (2), що важко зробити для n>2. Крім того, зі зростанням n може погіршуватися стійкість моделі (2). У випадку n=2 модель (2) добре вловлює зміну темпу розвитку пухлини та дозволяє фіксувати критичні точки (наприклад, момент переходу до плато чи вплив протипухлинної терапії). Поліноміальна регресія стала стандартним інструментом для аналізу даних у біостатистиці й часто слугує проміжною ланкою між простими та складнішими моделями (логістичною та Гомпертца). Зокрема, у змішаних моделях вона дозволяє моделювати як середню, так і індивідуальні траєкторії зростання [35].

2.3 Експоненціальна модель (модель Мальтуса)

Ця модель має прозору біологічну інтерпретацію, а саме швидкість зростання пропорційна масі (чисельності клітин) пухлини у даний момент часу, що відповідає біологічній тканині, усі клітини якої мають однакову ймовірність поділу, а зовнішні обмеження (ресурси, простір, імунна відповідь організму) ще не впливають на динаміку зростання. Історично ця модель була вперше запропонована англійським економістом Т. Мальтусом для зростання чисельності населення.

Математично експоненціальне зростання описується рівнянням:

$$\frac{dV}{dt} = bV + m, \quad V(t_0) = V_0, \tag{3}$$

де b – параметр Мальтуса, m – міграція клітин (переміщення від/до пухлини), розв'язок якого має вигляд $V(t_0) = (V_0 + m/b)e^{b(t-t_0)} - m/b$.

Таким чином, ця модель може використовуватися коли регресійний аналіз дає найкраще співвідношення до експоненціальної лінії тренду у даних часових рядів $\{V_i\}_{i=1}^{n}$. Серед основних переваг експоненціальної моделі — простота аналітичної форми, мінімальна кількість параметрів і легкість оцінки агресивності пухлини [12]. Вона ідеально підходить для ранніх фаз росту або для даних in vitro, коли вплив обмежуючих факторів мінімальний [11]. У популяційних та змішаних моделях параметр k дозволяє оцінити міжіндивідуальну варіабельність [35].

Втім, основне обмеження експоненціальної моделі — її непридатність для довгострокового прогнозування: у реальних біологічних системах ріст пухлин поступово уповільнюється через дефіцит кисню, поживних речовин та інші фактори [12,36]. В більшості випадків експоненціальна динаміка зберігається лише короткий час. У сучасних аналітичних системах експоненційну модель використовують як базову для порівняння складніших підходів або для попередньої оцінки параметрів.

2.4. Експоненціальна модель, яка переходить у лінійну

У ряді робіт розглядається комбінування експоненціальної та лінійної динаміки для адаптації до особливостей реальних експериментальних даних, у тому числі зниження темпів зростання завдяки використанню хіміотерапії [34, 12, 38]. У цих моделях пухлина зростає експоненціальне поки не досягне деякого критичного розміру $V = V^*$, після чого її зростання уповільнюється і апроксимується лінійною регресією, тобто

$$\frac{dV}{dt} = F(V) \equiv \begin{cases} \lambda V^* V, & V \le V^* \\ \lambda, & V > V^* \end{cases}, \qquad V(t_0) = V_0. \tag{4}$$

Відповідно до (4), перехід від експоненціального до лінійного зростання відбудеться у момент часу $t^* = t_0 + \left(\lambda V^*\right)^{-1} \ln\left(\lambda V^* / V\right)$.

Для обчислювальних цілей зручніше використовувати одне рівняння, яке описує перехід між фазами, особливо при врахуванні ефектів протипухлинних препаратів, що будуть введені в подальших дослідженнях. Тому в якості наближення використовується функція $F(w) = \lambda V^* V\left(1 + \left(V/V^*\right)^{\kappa}\right)^{-1/\kappa}$, де параметр κ визначає характер переходу між різними типами зростання. Ця функція узагальнює (3) і дозволяє моделювати плавний перехід між фазами зростання пухлини, зберігаючи можливість безперервного аналізу ефективності терапевтичних дій. Така модель добре відповідає зміні швидкості зростання різних типів пухлин під дією вдало підібраної терапії [34].

2.5. Логістична модель Ферхюльста.

Логістична модель стала однією з найпопулярніших у біостатистиці, оскільки дозволяє враховувати обмеження ресурсів та ефекти насичення, які властиві реальним біологічним системам [35, 36]. З часом біологічним клітинам бракує поживних речовин і кисню, наростає гіпоксія та метаболічні обмеження, які уповільнюють зростання тканини.

$$\frac{dV}{dt} = bV\left(1 - \frac{V}{V_{max}}\right), \quad V(t_0) = V_0, \tag{5}$$

де V_{max} - максимальний можливий розмір (маса) пухлини, b – швидкість зростання.

Розв'язок (5) із заданою початковою умовою має вигляд S-подібної кривої

$$V(t) = \frac{V_{max} V_0}{V_0 + (V_{max} - V_0)e^{-bt}}, \tag{6}$$

яка описує як швидку початкову фазу зростання пухлини, так і період повільного зростання з виходом на плато.

Перевага логістичної моделі – у можливості роботи з індивідуальними та популяційними даними (змішані ефекти, врахування гетерогенності [35, 36]). Водночас модель не враховує складних біологічних механізмів (наприклад, гетерогенність та некроз пухлини, імунний вплив) і в деяких випадках поступається за точністю моделі Гомпертца [36]. Логістична модель забезпечує оптимальний баланс між простотою та точністю і залишається універсальним інструментом для математичного опису зростання пухлин у біостатистиці.

Важлива модифікація моделі (5) зв'язана з урахуванням ефекту Алле [10] і має вигляд [39]

$$\frac{dV}{dt} = bV\left(1-\frac{V}{V_{max}}\right)\left(1-\frac{V}{V_{cr}}\right), \quad V(t_0) = V_0, \tag{7}$$

де $0 < V_{cr} < V_{max}$ - критичне значення, яке відповідає «переключенню» ракових клітин на «режим міграції» після досягнення критичного розміру пухлини.

Розв'язок (7) має вигляд

$$V(t) = \frac{V_{max}(V_0 - V_{cr}) + V_{cr}(V_{max} - V_0)\exp(bt(1 - V_{cr}/V_{max}))}{V_0 - V_{cr} + (V_{max} - V_0)\exp(bt(1 - V_{cr}/V_{max}))}, \tag{8}$$

причому функція $V(t)$ зростає при $V \in]V_{cr}, V_{max}[$ і зменшується при $V \in]0, V_{cr}[$.

Узагальнення моделі (5) на степеневу функцію у правій частині рівняння (узагальнена модель Річардса)

$$\frac{dV}{dt} = bV^{\alpha}\left(1-\left(\frac{V}{V_{max}}\right)^{\beta}\right)^{\gamma}, \quad V(t_0) = V_0, \tag{9}$$

в залежності від параметрів α і β має вигляд від лінійного до S-подібних розв'язків різної крутизни, які відносяться до класу функцій Річардса.

Наявність у моделі (9) чотирьох параметрів дає можливість апроксимації експериментальних кривих з достатньо високою точністю [45]. При $\alpha = \beta = \gamma = 1$ (9) переходить у класичну логістичну модель (5); при $\alpha = \gamma = 1, \beta \neq 1$ отримуємо модель Річардса; при $\alpha = 2/3, \beta = 1/3, \gamma = 1$ отримуємо логістичну модель у формі, яка була запропонована J.D. Murray [40] виходячи із загального закону збереження маси коли потік речовин до пухлини пропорційний площині її поверхні.

Оскільки клітинні цикли мають певну тривалість, синтез біохімічних компонентів, проліферація, міграція клітин та інші процеси впливають на подальші процеси зростання з деяким запізненням у часі τ, замість (5) розглядається рівняння із запізненням у часі [41, 42]

$$\frac{dV(t)}{dt} = bV(t-\tau)\left(1-\frac{V(t-\tau)}{V_{max}}\right). \tag{10}$$

2.6. Функція Берталанфі

Модель відповідає припущенню, що на клітинному рівні біологічне зростання є результатом балансу між анаболізмом і катаболізмом, тобто між синтезом і розпадом біологічних молекул, швидкості яких є степеневими функціями маси клітин. Таким чином, балансове рівняння зростання має вигляд:

$$\frac{dV}{dt} = bV^n - aV^m, \quad V(0) = V_0, \tag{11}$$

де a, b, n, m – параметри моделі, які залежать від використання ресурсів, метаболізму та морфологічної структури пухлини.

Розв'язок (11) у загальному вигляді є $w(t) = w_0 + \int_0^t \frac{dt}{bw^n - aw^m}$. Біологічно значущим є випадок $n = 1, m \leq 1$, для якого розв'язок можна обчислити шляхом інтегрування у вигляді

$V(t)=\left(\frac{a}{b}+\left(V_0^{1-m}-\frac{a}{b}\right)e^{b(1-m)t}\right)^{\frac{1}{1-m}}$. Наявність чотирьох параметрів у моделі (5) дає можливість гнучко апроксимувати експериментальні дані функцією Берталанфі, яка дає найкращі результати у випадку пухлин різних типів [43, 44, 45]. Також, по аналогії з (10), розглядаються рівняння Берталанфі (11) із запізненнями у часі.

2.7. Модель Гомпертца

Ця універсальна модель використовується для моделювання зростання чисельності популяцій мікроорганізмів, розмірів тіла тварин та окремих частин рослин, у тому числі пухлин, для аналізу динаміки смертності у демографічних дослідженнях та у моделюванні економічних процесів. Диференціальне рівняння моделі із початковою умовою

$$\frac{dV}{dt}=bVe^{-ct},\quad V(0)=V_0 \tag{12}$$

має розв'язок $V(t)=V_0\exp\left(\frac{b}{c}(1-\exp(-ct))\right)$, $b>0, 0<c<1$, де c – швидкість затухання проліферації.

Функція Гомпертца відтворює поступове уповільнення росту до появи насичення, а точка перегину функції розташована несиметрично, чим відрізняється від логістичної моделі [36]. Саме це дозволяє якісно моделювати дані для середніх і великих пухлин та уникати переоцінки об'єму у довгостроковому прогнозі [12]. З біологічної точки зору, гальмування зростання пов'язують з виснаженням ресурсів мікрооточення, наростанням гіпоксії, метаболічними зрушеннями і реакцією імунної системи організму [7]. Порівняльні дослідження показують, що саме крива Гомпертца для багатьох пухлин (гліоми, карциноми, раку грудей, моделі *in vivo*) часто дає найкраще узгодження з реальними траєкторіями зростання, особливо для тривалих періодів спостереження [11, 36, 46, 25].

Усі обговорені вище моделі (1) - (5), (7), (9), (11), (12) задовольняють отриманим експериментально ростовим кривим V(t) за рахунок підбору параметрів [10] і не існує однієї «найкращої» моделі [47]. Різні типи пухлин відповідають різним моделям, а також існують індивідуальні параметри, які пов'язані з особливостями організму пацієнта. Така невизначеність зворотних залежностей між динамікою зростання та конкретною математичною моделлю стимулює застосування сучасних алгоритмів, які використовують машинне навчання (Machine Learning, ML) і, особливо, глибоке машинне навчання (Deep Machine Learning, DML).

2.8. Статистичні методи аналізу даних

Статистичні методи аналізу даних спостережень та експериментальних вимірювань динаміки зростання пухлин у вигляді часових рядів широко використовуються для перевірки наявності закономірностей, кореляцій і т. д. для вибору однієї з математичних моделей (1)-(12), або їх комбінацій чи модифікацій для прогнозування динаміки онкологічної хвороби та її лікування. Для цього використовуються методи дескриптивної (описової) статистики, регресійний, кореляційний і факторний аналіз [47, 49]. Попередньо часові ряди мають бути згладжені (фільтрація випадкового шуму), перевірені на наявність пропущених даних та викидів, які пов'язані із недосконалістю процесу вимірювання. Як відомо, «сирі» медичні дані (raw data) рідко відповідають вимогам математичного моделювання. Вони можуть містити пропуски, аномалії, мати різні одиниці вимірювання, широкий діапазон варіацій та інші особливості, що суттєво ускладнює аналіз. Якісна попередня обробка сирих даних є необхідною умовою для подальшого застосування математичних моделей [35, 38].

2.8.1. Нормалізація даних

Нормалізація даних з приведенням до безрозмірних величин шляхом ділення на максимальне значення вимірюваної величини у часовому ряді, а також перевірка нормального (Гаусового) розподілу є важливою для застосування класичних статистичних методів. В результаті отримуємо часові ряди зі значеннями вимірюваних параметрів у діапазоні $V_i \in]0,1]$.

2.8.2. Викиди.

Медичні дані часто містять нетипові або помилкові значення, які можуть суттєво вплинути на результати моделювання. Для їх виявлення використовують як класичні статистичні методи -

аналіз boxplot, Z-score, IQR (інтерквартильний розмах), так і алгоритми ML. Викиди або виправляють, або виключають із вибірки залежно від контексту [11, 38].

2.8.3. Трансформації даних.

Для наближення розподілу до нормального часто застосовують логарифмічне, кореневе чи Box-Cox перетворення. Такі операції дозволяють підвищити ефективність регресійних моделей, особливо у випадках із сильними перекосами розподілу даних відносно Гаусового [35].

2.8.4. Імпутація пропусків.

У медичних наборах даних пропуски трапляються часто (наприклад, через нерегулярні обстеження, людський фактор або технічні збої). Для їх заповнення використовують прості підходи (середнє, медіана, останнє спостереження) або сучасні алгоритми, такі як метод k-найближчих сусідей (k-nearest neighbors, KNN), випадковий ліс (Random Forest, RF) мультиімпутація та інші [35, 38]. Це дозволяє уникнути втрати цінної інформації та підвищити статистичну достовірність аналізу.

2.8.5. Зниження розмірності.

При роботі з великим числом ознак (наприклад, різних характеристик з медичних зображень пухлин) актуальними є методи відбору (feature selection) та зниження розмірності датасету за допомогою аналізу головних компонент (Principal Component Analysis, PCA), алгоритму розподіленого стохастичного вбудовування сусідів (t-distributed stochastic neighbor embedding, t-SNE), та ряду інших. Це дозволяє зосередитися на найбільш інформативних параметрах та уникнути перенавчання.

Комплексне застосування зазначених статистичних процедур забезпечує максимальну якість і достовірність аналізу, дозволяє працювати з великими, гетерогенними, частково неповними наборами даних — тобто в умовах, які характерні для сучасної біоінформатики та клінічної онкології [12, 38].

2.8.6. Регресійний і кореляційний аналіз.

Різни типи ліній регресії (лінійна, поліноміальна, степенева, експоненціальна, логарифмічна, S-подібна) будувалися методом найменших квадратів [49]. Найкраща значуща регресія обралась відповідно до максимального значення коефіцієнта детермінації

$$R^2 = 1 - \sum_{i=1}^{n} \left(V_i - V_i^{mod} \right)^2 / \sum_{i=1}^{n} \left(V_i - \bar{V}_i \right)^2 , \tag{13}$$

де $\{V_i\}_{i=1}^n$ - виміряні часові ряди, $\bar{V}_i$ - середнє значення ряду, V_i^{mod} - значення, прогнозовані за обраною математичною моделлю.

У залежності від отриманої регресійної кривої, обирається відповідна математична модель (лінійного, поліноміального, експоненціального та ін. зростання).

Кореляції між двома кривими зростання пухлин і обчислювалась на основі коефіцієнта кореляції Пірсона

$$K_P = \frac{\sum_{i=1}^{n} (X_i - M(X)(Y_i - M(Y))}{\sqrt{\sum_{i=1}^{n} (X_i - M(X))^2 (Y_i - M(Y))^2}} , \tag{14}$$

де $\{X_i\}_{i=1}^n$ і $\{Y_i\}_{i=1}^n$ - часові ряди, $M(X_i)$, $M(Y_i)$ - їх середні значення (математичні очікування).

2.8.7. Критерії якості обраної математичної моделі.

Для порівняння різних масивів даних потрібні обчислення середніх значень (математичне очікування, медіана, мода) та розкиду даних (дисперсія або стандартне відхилення). Для перевірки точності математичної моделі найчастіше використовують критерій R^2 (13). Статистично значущими є значення $R^2 > 0.8$, але у біомедичних даних часто використовують невеликі датасети даних, які отримані на групах із 20-40 хворих, і тому статистичні методи, строго кажучи, не можуть бути використані, а значення $R^2 > 0.5$ вважаються за істотними.

Для кількісної оцінки точності тієї чи іншої математичної моделі використовуються також середня абсолютна помилка (Mean Absolute Error, MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|V_i - V_i^{mod}\right|, \quad (15)$$

або середньоквадратична помилка (Root Mean Squared Error, RMSE)

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}\left(V_i - V_i^{mod}\right)^2\right)^{1/2}, \quad (16)$$

які чутливі до великих відхилень і добре підходять для датасетів із рідкими, але значними аномаліями [12].

Крім основних статистичних показників (13), (15), (16), для порівняння математичних моделей із різною складністю використовуються інформаційні критерії, такі як Akaike Information Criterion (AIC) та Bayesian Information Criterion (BIC). Вони дозволяють визначати, чи виправдане збільшення числа параметрів поліпшенням точності, і допомагають уникати перенавчання алгоритмів ML. Додатково для оцінки нормальності розподілу залишків може застосовуватись тест Колмогорова-Смірнова чи аналіз автокореляції залишків тестом Дарбіна-Вотсона [12, 36].

Завдяки цим інструментам стає можливим обирати не лише модель із найкращою підгонкою, а й ту, що має оптимальний баланс між точністю, складністю і стійкістю до шуму у даних. Таким чином, стандартні статистичні методи потребують ретельної підготовки сирих даних до аналізу, що займає значний час. При цьому сучасні методи ML, DML можуть ефективно обробляти сирі дані, що значно зберігає час.

2.9. Підхід Random Forest (RF)

RF - це ансамблевий метод машинного навчання, який дозволяє ефективно моделювати складні нелінійні залежності у біомедичних задачах, зокрема для аналізу динаміки зростання пухлин [38, 50]. Головна ідея методу - це побудова ансамблю дерев рішень, де кожне дерево тренується на випадковий підвибірці даних і підмножині ознак, а прогноз визначається як середнє усіх дерев. Така архітектура забезпечує стійкість до викидів, автоматичну обробку пропусків та зниження ризику перенавчання моделі. Порівняно з класичними математичними моделями (1)-(12), RF не потребує визначення конкретної функціональної форми й може апроксимувати навіть складні S-подібні чи багатофазні криві зростання. Особливо актуально це для тих задач, у яких дані вимірювань містять шум, а структура залежностей між даними заздалегідь невідома [51].

Серед ключових переваг RF — гнучкість, висока точність прогнозування, можливість роботи з великою кількістю предикторів і ефективна інтеграція різних типів ознак (радіомічних, генетичних, клінічних). У клінічний онкології метод використовується для індивідуального прогнозування траєкторій зростання пухлини, аналізу факторів ризику та виживаності, а також для створення цифрових двійників пацієнтів (digital twin) [38, 50]. Недоліки RF — відсутність явної біологічної інтерпретації параметрів, складність аналізу впливу окремих змінних і потреба у достатньо великий вибірці для досягнення стабільної точності моделі. Особливий розвиток отримали Випадкові ліси виживання (Random Survival Forests) — модифікація RF для аналізу часу до певної події, що дуже корисно у медичних дослідженнях [50].

Таким чином, комбінація класичних статистичних методів та сучасних алгоритмів підготовки і обробки даних вимірювань дозволяють знайти найбільш адекватну математичну модель, провести її валідацію на датасетах медичної інформації та використовувати для прогнозу динаміки зростання пухлини та вибору персоніфікованого лікування. Математичні моделі (1)-(12) у вигляді одного чи системи ЗДР можуть бути розширені на неоднорідні пухлини з їх мікрооточенням шляхом додавання відповідних дифузійних членів у рівняння, що дає системи ДРЧП для більш детального опису просторово-часових процесів, які пов’язані із розвиненням пухлини та змінами оточуючих її тканин.

3. Матеріали і методи.

3.1. Джерело інформації.

У цій роботі виконаний порівняльний аналіз статистичних та сучасних машинних методів на прикладі гліобластоми (GBM), дані зростання якої отримано з набору даних LUMIERE (Longitudinal Glioblastoma MRI with expert RANO evaluation), яка міститься у відкритому джерелі [52].

База даних охоплює 91 пацієнта з діагнозом гліобластома, для яких проведено 638 МРТ-обстежень (загалом 2487 окремих зображень). Кожне обстеження включає мультипараметричні МРТ-зображення, отримані в основних режимах: T1 (до введення контрасту), T1c/T1Gd (після введення контрасту), T2w та FLAIR (Fluid-Attenuated Inversion Recovery). Така комбінація забезпечує детальну візуалізацію структури пухлини та прилеглих тканин. До набору додаються результати автоматичної сегментації пухлинних зон, які надаються разом із МРТ-зображеннями.

Тут виокремлюють три ключові компоненти пухлини:

1. Некротична/неактивна зона NEC (Necrotic/NonEnhancing core): центральна частина пухлини, яка часто складається з відмерлих клітин і не накопичує контрастну речовину.
2. Ділянка з контрастним підсиленням CEL (Enhancing Core/Contrast Enhancing Lesion): активна частина пухлини, що інтенсивно накопичує контраст і свідчить про агресивне зростання.
3. Зона набряку PEC (Edema/Peritumoral Edema Compartment): область навколо пухлини, що характеризується накопиченням рідини внаслідок запальних процесів або порушення відтоку.

У супровідних JSON-файлах (JavaScript Object Notation) подано об’єми кожного з компонентів пухлини, що дозволяє проводити аналіз її змін з часом. Додатково, у наборі є дані про пацієнтів: стать, вік, особливості гена MGMT, наявність мутації в гені IDH1, тривалість життя після лікування та результати терапії за шкалою RANO (Response Assessment in Neuro-Oncology). Для цілей даного дослідження було обрано підмножину з 599 випадків, у яких представлені всі три компоненти пухлини, що є репрезентативною вибіркою для статистично достовірного аналізу даних. Багатовимірний характер бази даних LUMIERE, яка поєднує зображення, кількісні параметри та клінічні характеристики, створює базу для комплексного аналізу динаміки гліобластоми з урахуванням індивідуальних особливостей пацієнтів.

3.2. Попередня обробка та структурування даних для аналізу.

Початкові дані були збережені у форматі JSON та містили числові показники об’ємів сегментів пухлини. З метою їх подальшого аналізу ці дані були імпортовані й упорядковані у формат табличних структур із використанням мови програмування Python та бібліотеки pandas. Основну увагу було приділено змінним, що характеризують об’єми трьох ключових компонентів пухлини: NEC, CEL, і PEC для всіх доступних моментів спостереження у часі, що вимірювався у тижнях.
У подальшому ці дані були об’єднані з інформацією із додаткових таблиць, що включали:

- Демографічну інформацію: вік в якому було проведено хірургічне втручання (Age at surgery (years)) та стать (Sex) пацієнтів;
- Клінічні змінні: статус метилування промотора гена MGMT (MGMT qualitative), мутаційний статус IDH (IDH (WT: wild type), IDH method);
- Категорії відповіді за критеріями RANO (Rating);
- Загальний час виживання (LessThan3Months як додаткова змінна).

Інтеграція здійснювалася шляхом злиття таблиць за спільними колонками за унікальним ідентифікатором пацієнта (Patient) та відповідною часовою точкою (Wheel). На наступному етапі було проведено очистку даних. З метою забезпечення коректності математичних операцій усі числові поля було приведено до відповідних типів (float, int), а категоріальні – до уніфікованих ярликів. Назви колонок було стандартизовано (видалено спеціальні символи, замінено пропуски на підкреслення тощо) для підвищення зручності роботи з даними.

У результаті попередньої обробки був сформований структурований дата-фрейм, який включає 599 записів (по одному на кожне спостереження), а також, залежно від наявності допоміжної інформації, 10–12 змінних, а саме:

1. Patient_N: Унікальний ідентифікатор пацієнта (N=1, 2,…,91).
2. Week: Інтервал часу у тижнях відносно першого дослідження.
3. NEC, CEL, PEC: Об’єми відповідних сегментів пухлини (мм^3, тип float).
4. Age (роки): Вік пацієнта на момент операції (тип int).
5. Sex: Стать пацієнта (‘male‘ або ‘female‘).
6. IDH (WT: wild type): Мутаційний статус IDH (‘Mutated‘, ‘WT‘, ‘Unknown‘).

7. IDH method: Метод визначення статусу IDH.
8. MGMT qualitative: Метилування промотора MGMT (якісна і кількісна оцінка).
9. Rating: Категорія відповіді за критеріями RANO (‘Complete Response‘, ‘Partial‘, ‘Stable‘, ‘Progressive‘).
10. LessThan3Months: Двійкова змінна, яка вказує на виживання менше 3 місяців після операції (0 або 1).
11. Date: Дата МРТ-дослідження.
12. NonMeasurableLesions: Додаткова характеристика уражень, що не підлягають вимірюванню.

У досліджуваній групі середній вік становив ~60,7±14,6 років. Розподіл за статтю був майже рівномірним: 44 жінки та 47 чоловіків. Статус метилування MGMT був відомий для 80 пацієнтів: 37 мали метильований промотор, 43 – неметильований, ще 11 – невідомий. Статус IDH1 був відомий для 58 пацієнтів, серед яких лише один мав виявлену мутацію (решта – WT). Через надто малу чисельність групи з IDH1-мутацією порівняльний аналіз за цим маркером не проводився.

3.3. Результати статистичного аналізу датасетів.

Для кожного компонента пухлини обчислено базові статистичні характеристики: середнє значення, медіана, стандартне відхилення та діапазон значень (Таблиця 1). Отримані результати свідчать про суттєву міжіндивідуальну варіативність об’ємів пухлин. Для виявлення потенційних залежностей між об’ємами компонент проведено кореляційний аналіз із використанням коефіцієнту Пірсона (14). Найвищу кореляцію виявлено між CEL і PEC ($R^2 = 0.54$), що свідчить про помірний взаємозв’язок між зростанням активної частини пухлини та перифокальним набряком. Кореляція між NEC та іншими компонентами є низькою (R^2 =0,14–0,33). Відповідна матриця кореляцій наведена на Рис.1.

Таблиця 1. Середні значення для різних компонентів гліобластоми (GBL)
Table 1. Average values for different components of glioblastoma (GBL)

Компонент	Діапазон (мм^3)	Середнє (мм^3)	Медіана (мм^3)	Дисперсія (мм^3)
NEC	0-150	25.0	15.0	30.0
CEL	0-300	40.8	30.0	35.0
PEC	5-450	62.7	45.0	50.0

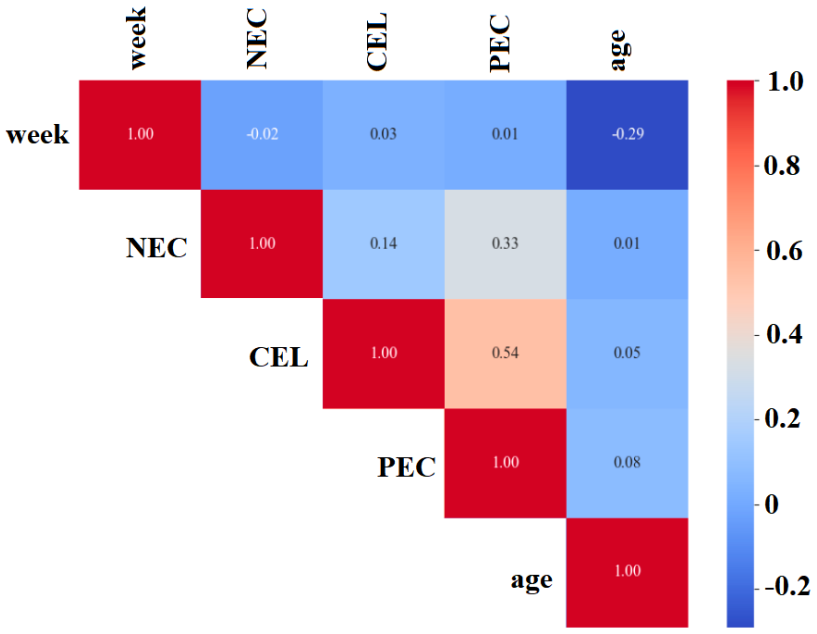


Рис.1. Кореляційна матриця числових ознак.
Fig. 1. Correlation matrix of numerical features.

Додатково проведено порівняння обсягів компонент між підгрупами пацієнтів, зокрема за статусом MGMT та статтю. Найбільш виражені відмінності спостерігались між групами MGMT: пацієнти з метильованим промотором мали статистично значущо менші значення Enhancing_Core порівняно з неметильованими ($p < 0{,}05$), що узгоджується з попередніми клінічними даними про кращий прогноз і менш агресивний ріст при MGMT-метиляції.

Щодо статі пацієнта, розподіл обсягів CEL між чоловіками та жінками був подібним: медіанні значення та інтерквартильний розмах практично не відрізнялись, а кількість викидів була приблизно однаковою в обох групах. Це узгоджується з літературними даними про відсутність істотного гендерного впливу на макроструктурні характеристики GBM [47].

Аналіз впливу віку на об'єм CEL продемонстрував, що різниця між молодшою (<60 років) та старшою (≥60 років) виковими групами статистично незначуща (U-критерій Манна-Уітні, $p = 0.787$). Незважаючи на певну тенденцію до вищих медіанних значень у старшій групі, результат не дозволяє стверджувати про суттєвий вплив віку на розвиток перифокального набряку у пацієнтів цієї групи.

Окрім порівняльного аналізу, було досліджено динаміку змін об'ємів пухлин у часі для окремих пацієнтів. У якості прикладу, на Рис.2 наведені часові ряди об'ємів трьох компонент пухлини для пацієнта Patient-025. Помітно поступове нелінійне зростання CEL та PEC. Середні об'єми компонент пухлин та їх розкид за даними 599 випадків наведені на Рис.3. Як видно з рисунка, компонент PEC характеризується найбільшими об'ємами та високою варіабельністю. Натомість об'єми NNE і CEL є меншими, хоча теж мають значну кількість викидів.

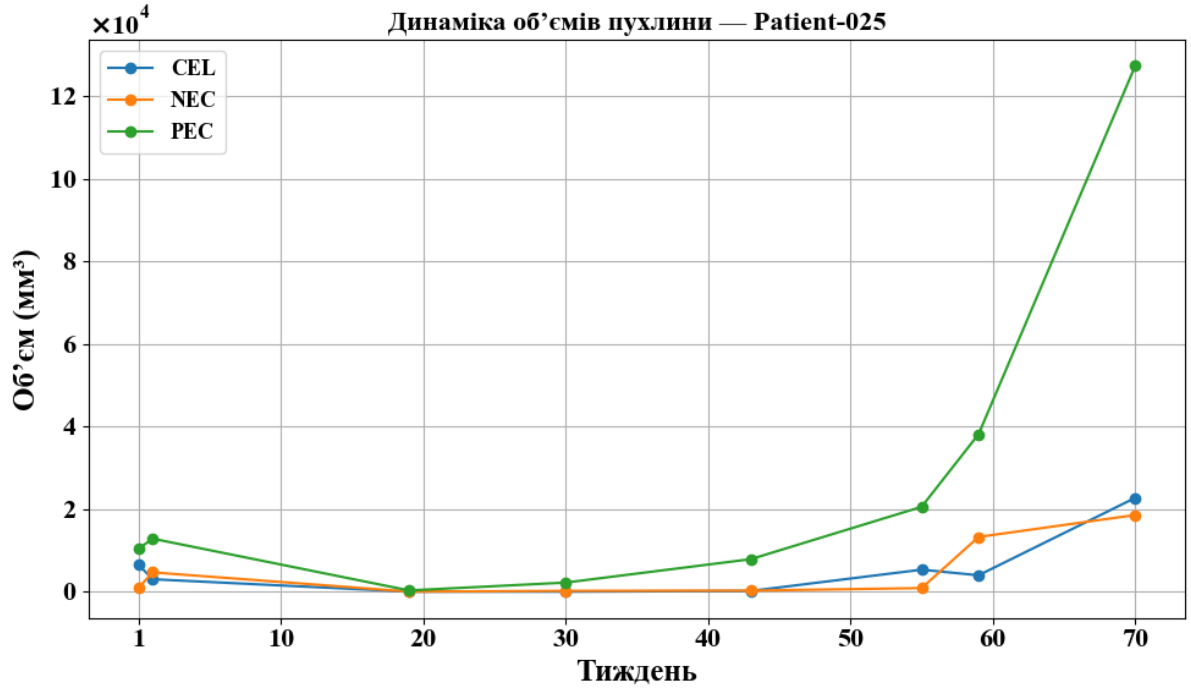


Рис.2. Динаміка у часі об'ємів NEC, CEL і PEC пухлини для пацієнта Patient-025.
Fig. 2. Dynamics of NEC, CEL, and PEC tumor volumes over time for Patient-025.

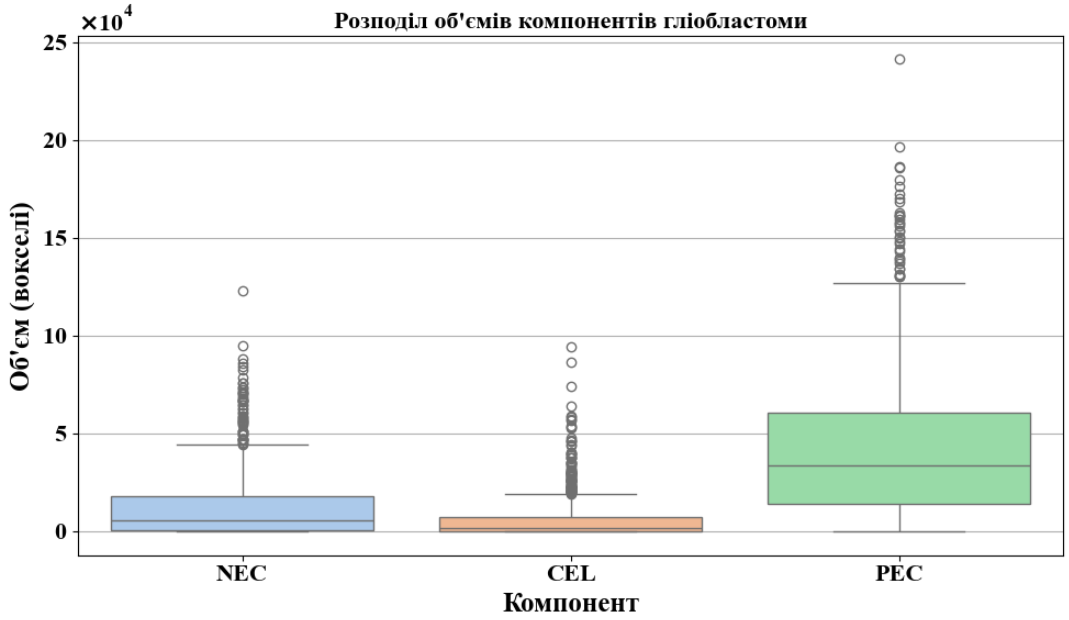


Рис.3. Розподіл об’ємів компонентів гліобластоми: NEC, CEL, PEC.
Fig. 3. Distribution of glioblastoma component volumes: NEC, CEL, PEC.

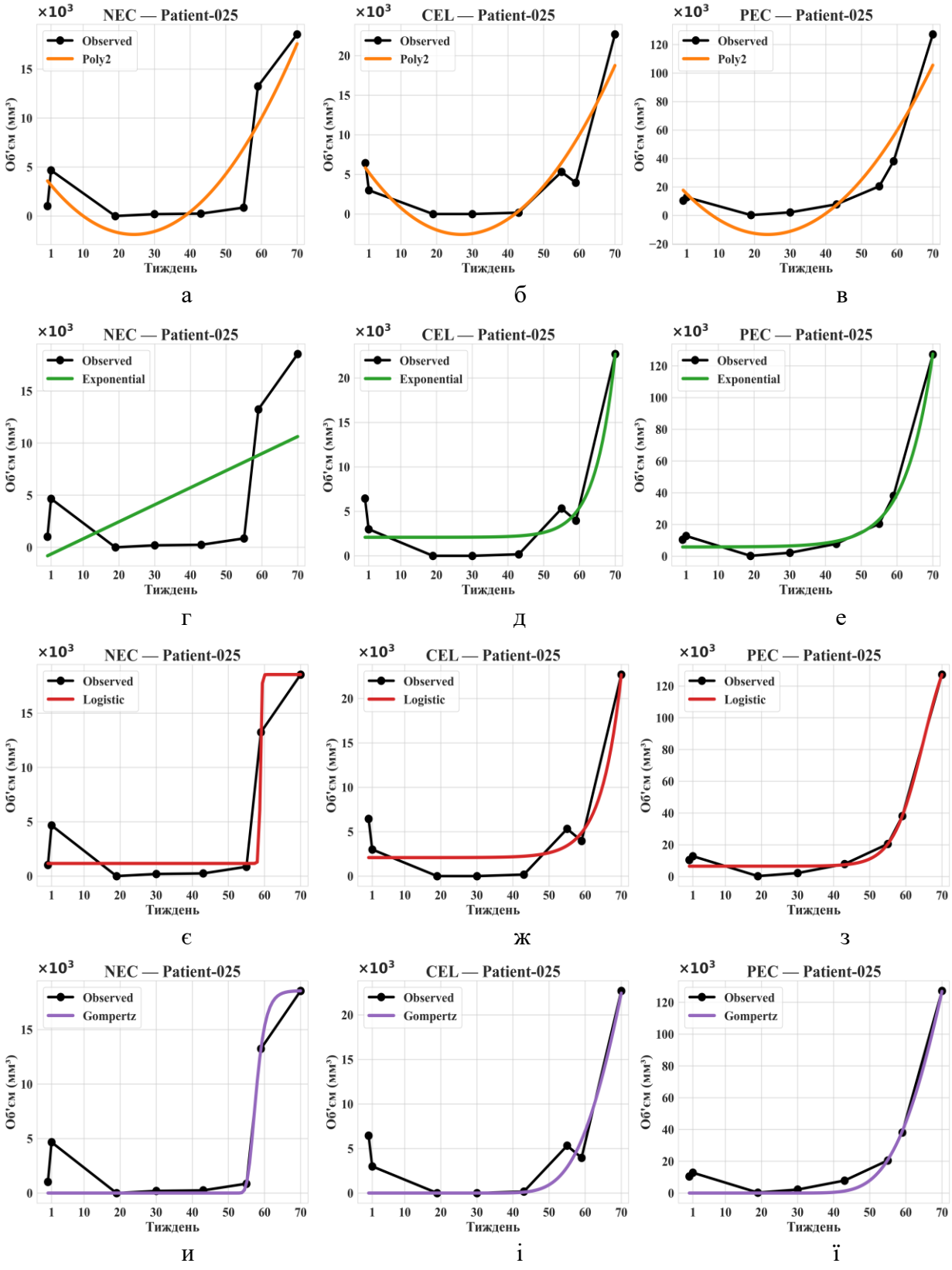
Загалом, результати описової статистики, порівняльного та кореляційного аналізу дозволяють сформувати цілісне уявлення про варіативність об’ємів компонент GBM у групі пацієнтів. Виявлені зв’язки (зокрема між об’ємами активного ядра та набряку) і тенденції (менші об’єми при MGMT-метиляції) можуть бути використані для параметризації математичних моделей росту пухлини. Додатково, наявність повних часових рядів створює передумови для побудови індивідуалізованих моделей динаміки зростання пухлини.

4. Математичне моделювання зростання пухлини.

4.1. Регресійний аналіз.

Регресійний аналіз індивідуальних кривих зростання пухлин пацієнтів показав відсутність лінійної регресії (R^2~0.1-0.3), більш значуще наближення поліномами 2-го порядку (R^2~0.37-0.72) та експонентною (R^2~0.224-0.623). Експоненціальна модель виявилася ефективною для окремих траєкторій із вираженим початковим зростанням, проте для тривалих або складних кривих її прогностичні властивості недостатні. Наближення S-подібними кривими показало значущі результати для логістичної кривої (R^2~0.371-0.690) і невисокі результати – для функції Гомпертца (R^2~0.012-0.029).

Метод RF з 200 деревами і обмеженням глибини (max_depth=3) був застосований для кривих зростання кожної з 3-х частин пухлини кожного з 599 хворих. Така конфігурація була обрана для запобігання перенавчанню, враховуючи обмежену кількість точок для кожного випадку. На кожному кроці фіксувалися значення метрик якості R^2, RMSE, MAE, MAPE а також параметри моделі. У якості приклада результати регресійного наближення кривих зростання об’ємів NEC, CEL, PEC у даних Patient-025 наведені на Рис.4. Середні значення якості наближень (значення R^2 та RMSE) наведені на Рис.5. Таким чином, динаміка зростання CEL характеризується помірно високими значеннями R^2~0.37 і мінімальним значенням RMSE. Криві зростання PEC мають найбільші значення R^2, але значно більший розкид. Динаміка зростання NEC слабо піддається використаним регресійним моделям за рахунок раптових змін у швидкості зростання з прискореннями та уповільненнями.

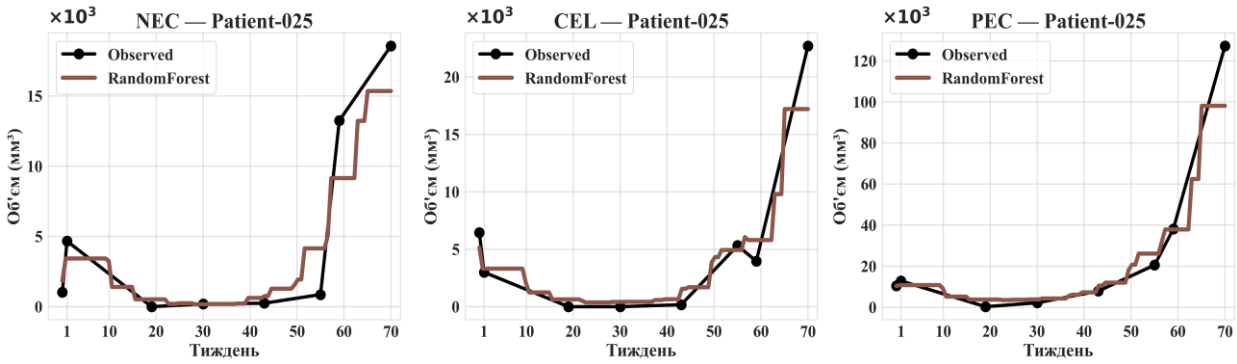


а б в

г д е

є ж з

и і ї



К Л М

Рис.4. Результати апроксимації результатів вимірювання кривих зростання об'ємів NEC (а,г,є,и,к), CEL (б,д,ж,і,л), PEC (в,е,з,ї,м) параболічною (а-в), експоненціальною (г-е), логістичною (є-з), гомперцевою (и-ї) та RF (к-м) функціями для пацієнта Patient-025.

Fig. 4. Results of approximation of measurement results for NEC (a, d, e, f, g) and CEL (b, c, d, e, f) growth curves, PEC (c, e, g, h, m) growth curves using parabolic (a-c), exponential (d-e), logistic (f-g), Gompertz (h-i) and RF (k-m) functions for patient Patient-025.

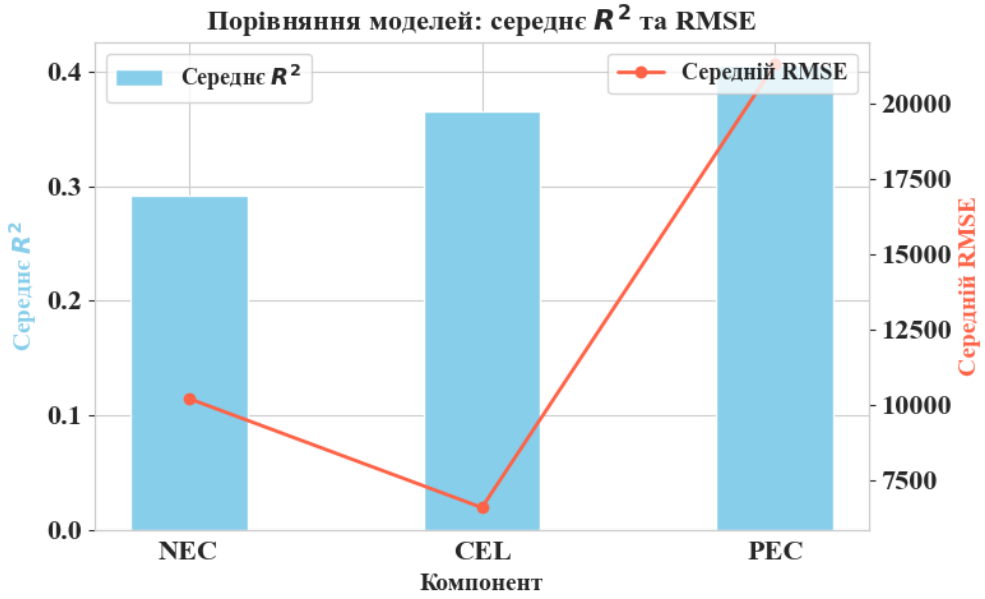


Рис.5. Порівняння середніх значень R2 та RMSE для кривих зростання NEC, CEL, PEC для Patient-025.

Fig. 5. Comparison of average R2 and RMSE values for NEC, CEL, and PEC growth curves for Patient-025.

4.2. Прогнозування динаміки зростання пухлин.

Після визначення найкращого (у сенсі параметрів R^2, MAE, RMSE та ін.) наближення (регресії) до кривої зростання відповідної області пухлини (NEC, CEL, PEC) у даного пацієнта, можна прогнозувати динаміку зростання шляхом інтерполяції. У роботі було здійснено прогноз об'єму пухлини на 8 тижнів наперед, що відповідає типовому проміжку між контрольними обстеженнями у клінічний практиці. На основі моделі найкращого наближення можна провести екстраполяцію кривої від останніх точок (59 та 70 тижнів) до наступної (78 тижнів) відповідно до моделі. На Рис.6 наведений приклад прогнозування для однієї з кривих. Крім того, можна провести ідентифікацію параметрів відповідної математичної моделі (2)-(5), (7), (9), (11) або (12) та використовувати її розв'язок із відомими коефіцієнтами для прогнозування як подальшого зростання пухлини, так і можливих наслідків її лікування різними препаратами. Для цього потрібні лабораторні дані про чутливість даної пухлини до ліків різних типів. Відповідний доданок треба ввести до правої частини ЗДР у якості негативного джерела для V(t) у вигляді лінійної або періодичної функції від дози препарату з урахуванням поглинання та розкладу препарату в організмі з часом, відповідно до кінетики Міхаеліс-Менетен або іншої.

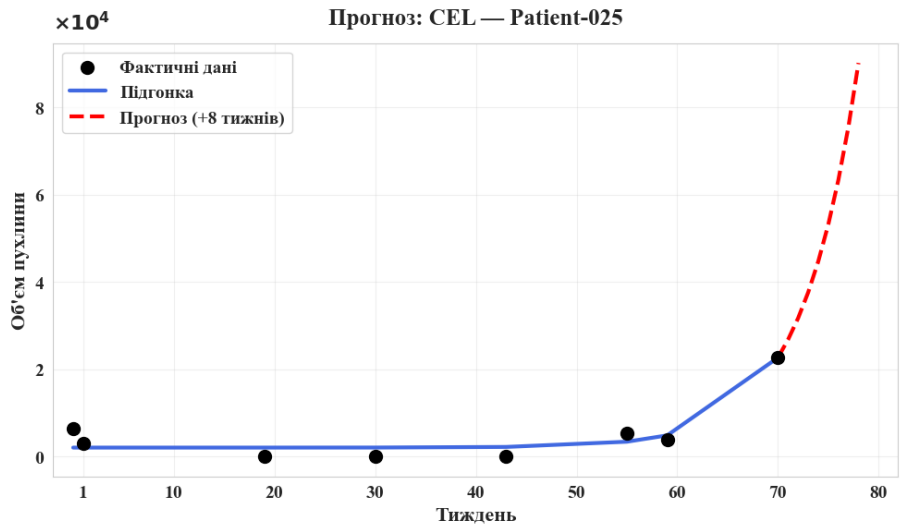


Рис.6. Прогноз динаміки об'єму компоненти CEL у пацієнта Patient-025.
Fig. 6. Forecast of the dynamics of the CEL component volume in patient Patient-025.

6. Висновки.

Результати роботи підтвердили практичну цінність математичного моделювання для оцінки динаміки зростання пухлин, можливостей оптимізації терапії та прогнозування результатів лікування. Поєднання аналітичних моделей із алгоритмами ML, разом із якісною підготовкою даних (нормалізація, імпутація пропусків, згладження, перетворення ознак) відкриває перспективи для подальшого розвитку персоналізованих підходів у сучасній медичний практиці та біостатистиці.

На основі аналізу одного клінічного випадку проведене порівняння класичних математичних моделей зростання пухлини (лінійна, поліноміальна, експоненційна, логістична, Берталанфі та Гомпертца), а також сучасних інструментів ML, зокрема RF. Кожна з моделей продемонструвала свої сильні та слабкі сторони: лінійна і експоненціальна моделі добре описують початкові фази, логістична та Гомпертцева — динаміку із фазою насичення, а RF забезпечує адаптацію до складних, гетерогенних і частково неповних даних, однак має слабку інтерпретованість.

Виявлено, що для тривалих та насичених спостережень найкраще працюють нелінійні моделі, тоді як короткі часові ряди доцільно аналізувати за допомогою лінійної чи поліноміальної регресії, для яких є прозора біологічна інтерпретація. Використання сучасних алгоритмів ML значно розширює можливості кількісного аналізу пухлинного росту, особливо у випадках складної структури даних чи відсутності чіткої аналітичної форми моделі.

Саме комбінований підхід - об'єднання класичних та сучасних статистичних інструментів - дає змогу отримати гнучкі та інтерпретовані рішення для персоналізованої медицини. Результати можуть стати основою для впровадження цифрових двійників (digital twin) пацієнта, підвищення точності діагностики і розробки нових індивідуалізованих стратегій у клінічний онкології.

REFERENCES

1. World Health Organization, Cancer, 3 February 2025, [Online] Available: https://www.who.int/news-room/fact-sheets/detail/cancer Accessed on: February 01, 2026.
2. R.L. Siegel, T.B. Kratzer, A.N. Giaquinto, et al., "Cancer statistics, 2025", CA: A Cancer Journal for Clinicians, 75(1), 10–45, 2025. doi:10.3322/caac.21871
3. A.O. Lawal, T.J. Ogunniyi, O.I. Oludele, et al., "Innovative laboratory techniques shaping cancer diagnosis and treatment in developing countries", Discover Oncology, 16, 137, 2025. doi:10.1007/s12672-025-01877-w
4. G. Molla, M. Bitew, "The Future of Cancer Diagnosis and Treatment: Unlocking the Power of Biomarkers and Personalized Molecular-Targeted Therapies", Journal of Molecular Pathology, 6, 20, 2025. doi:10.3390/jmp6030020

5. H.K. Elaibi, F.F. Mutlag, E. Halvaci, et al., “Comparison of Traditional and Modern Diagnostic Methods in Breast Cancer”, Measurement, 242, 116258, 2025. doi:10.1016/j.measurement.2024.116258
6. R.P. Araujo, D.L.S. McElwain, “A history of the study of solid tumour growth: the contribution of mathematical modelling”, Bulletin of Mathematical Biology, 66(5), 1039–1091, 2004. doi: 10.1016/j.bulm.2004.03.001
7. D. Hanahan, R.A. Weinberg, “The Hallmarks of Cancer”, Cell, 100(1), 57–70, 2000. doi:10.1016/S0092-8674(00)81683-9
8. J. Folkman, “Tumor Angiogenesis: Therapeutic Implications”, The New England Journal of Medicine, 285, 1182–1186, 1971. doi:10.1056/NEJM197111182852108
9. P. Carmeliet, R.K. Jain, “Angiogenesis in Cancer and Other Diseases”, Nature, 407(6801), 249–257, 2000. . doi:10.1038/35025220
10. W.C. Allee, E. Bowen, “Studies in Animal Aggregations: Mass Protection Against Colloidal Silver Among Goldfishes”, Journal of Experimental Zoology, 61(2), 185–207, 1932. doi:10.1002/jez.1400610202
11. E.A. Sarapata, L.G. de Pillis, “A Comparison and Catalog of Intrinsic Tumor Growth Models”, Bulletin of Mathematical Biology, 76(8), 2010–2024, 2014. doi:10.1007/s11538-014-9986-y
12. S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J.M.L. Ebos, et al., “Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth”, PLoS Computational Biology, 10(8), e1003800, 2014. doi:10.1371/journal.pcbi.1003800Laird, A. Dynamics of Tumour Growth: Comparison of Growth Rates and Extrapolation of Growth Curve to One Cell. Br J Cancer 19, 278–291 (1965). doi:10.1038/bjc.1965.32
13. A. Laird, “Dynamics of Tumour Growth: Comparison of Growth Rates and Extrapolation of Growth Curve to One Cell.”, Br J Cancer 19, 278–291 (1965). doi:10.1038/bjc.1965.32
14. L. von Bertalanffy, “Quantitative laws in metabolism and growth”, The Quarterly Review of Biology, 32(3), 217–231, 1957. doi: 10.1086/401873
15. P. Armitage, R. Doll, “The age distribution of cancer and a multi-stage theory of carcinogenesis”, British Journal of Cancer, 8(1), 1–12, 1954. doi: 10.1038/bjc.1954.1
16. A.C. Burton, “Rate of growth of solid tumours as a problem of diffusion”, Growth, 30(2), 157–176, 1966.
17. H.P. Greenspan, “Models for the growth of a solid tumor by diffusion”, Studies in Applied Mathematics, 51, 317–340, 1972. doi: 10.1002/sapm1972514317
18. T. Teorell, “Kinetics of distribution of substances administered to the body. I”, Archives Internationales de Pharmacodynamie et de Thérapie, 57, 205–225, 1937.
19. T. Teorell, “Kinetics of distribution of substances administered to the body. II”, Archives Internationales de Pharmacodynamie et de Thérapie, 57, 226–240, 1937.
20. R.E. Bellman, J.A. Jacquez, R. Kalaba, “Mathematical models of chemotherapy”, in: J. Neyman (ed.), Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Cambridge University Press, 57–65, 1961.
21. R.C. Rockne, J.G. Scott, “Introduction to Mathematical Oncology”, JCO Clinical Cancer Informatics, 3, 1–4, 2019. doi: 10.1200/CCI.19.00010
22. D. Wodarz, N. Komarova, Dynamics of Cancer: Mathematical Foundations of Oncology, World Scientific, 2014. doi: 10.1142/8846
23. N. Kizilova, “Mathematical modelling of biological growth and tissue engineering”, in: R. Bedzinski, M. Petrtyl (eds.), Current trends in development of implantable tissue structures, Lecture Notes of the ICB Seminar, Warsaw, 18–27, 2012.
24. N.N. Kizilova, S.A. Logvenkov, A.A. Stein, “Mathematical modeling of transport-growth processes in multiphase biological continua”, Fluid Dynamics, 47(1), 1–9, 2012. doi: 10.1134/S0015462812010012
25. F. Kozusko, Ž. Bajzer, “Gompertzian growth and cell population dynamics”, Mathematical Biosciences, 185, 153–167, 2003. doi: 10.1016/S0025-5564(03)00086-2
26. S. Xu, “Analysis of a delayed mathematical model for tumor growth”, Nonlinear Analysis: Real World Applications, 11, 4121–4127, 2010. doi: 10.1016/j.nonrwa.2010.04.001
27. H. Byrne, “The effect of time delays on the dynamics of avascular tumor growth”, Mathematical Biosciences, 144, 83–117, 1997. doi: 10.1016/S0025-5564(97)00034-0

28. M. Kamran, J.Y. Abdullah, A.S. Ahmad Satmi, M. Genisa, A. Majeed, T. Nadeem, “Mathematical modeling and analysis of tumor growth models integrating treatment therapy”, Mathematics and Computers in Applications, 30, 119, 2025. doi: 10.3390/mca30060119
29. H. Byrne, M.A.J. Chaplain, “Growth of necrotic tumors in the presence and absence of inhibitors”, Mathematical Biosciences, 135, 187–216, 1996. doi: 10.1016/0025-5564(96)00011-5
30. H.P. Greenspan, “On the growth and stability of cell cultures and solid tumors”, Journal of Theoretical Biology, 56, 229–242, 1976. doi: 10.1016/S0022-5193(76)80054-9
31. F.J. Solis, S.E. Delgadillo, “Discrete mathematical models of an aggressive heterogeneous tumor growth with chemotherapy treatment”, Mathematical and Computer Modelling, 50, 646–652, 2009. doi: 10.1016/j.mcm.2009.05.010
32. A. Qi, X. Zheng, C. Du, B. An, “A cellular automaton model of cancerous growth”, Journal of Theoretical Biology, 161, 1–12, 1993. doi: 10.1006/jtbi.1993.1031
33. R.A. Gatenby, R.J. Gillies, “A microenvironmental model of carcinogenesis”, Nature Reviews Cancer, 8(1), 56–61, 2008. doi: 10.1038/nrc2255
34. P. Magni, M. Simeoni, I. Poggesi, M. Rocchetti, G. De Nicolao, “A mathematical model to study the effects of drug administration on tumor growth dynamics”, Mathematical Biosciences, 200, 127–151, 2006. doi: 10.1016/j.mbs.2005.12.028
35. B. Ribba, N.H.G. Holford, P. Magni, et al., “A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis”, CPT: Pharmacometrics & Systems Pharmacology, 3(5), e113, 2014. doi: 10.1038/psp.2014.14
36. C. Vaghi, A. Rodallec, R. Fanciullino, et al., “Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors”, PLoS Computational Biology, 16(2), e1007178, 2020. doi: 10.1371/journal.pcbi.1007178
37. H.G. Müller, U. Stadtmüller, “Generalized functional linear models”, The Annals of Statistics, 33(2), 774–805, 2005. doi: 10.1214/009053604000001156
38. G. Lorenzo, S.R. Ahmed, D.A. Hormuth, et al., “Patient-specific, mechanistic models of tumor growth incorporating artificial intelligence and big data”, Annual Review of Biomedical Engineering, 26(1), 529–560, 2024. doi: 10.1146/annurev-bioeng-081623-025834
39. K. Böttger, H. Hatzikirou, A. Voss-Böhme, E.A. Cavalcanti-Adam, M.A. Herrero, A. Deutsch, “An emerging Allee effect is critical for tumor initiation and persistence”, PLoS Computational Biology, 11(9), e1004366, 2015. doi: 10.1371/journal.pcbi.1004366
40. J.D. Murray, Mathematical Biology: An Introduction, Springer, New York, 2002, pp. 1–75. doi: 10.1007/b98868
41. R. Schuster, H. Schuster, “Reconstruction models for the Ehrlich ascites tumor of the mouse”, in: O. Arino, D. Axelrod, M. Kimmel (eds.), Mathematical Population Dynamics, Vol. 2, Wuertz, Winnipeg, Canada, 335–348, 1995.
42. U. Foryś, A. Marciniak-Czochra, “Logistic equations in tumour growth modelling”, International Journal of Applied Mathematics and Computer Science, 13(3), 317–325, 2003.
43. H.H. Diebner, T. Zerjatke, M. Griehl, I. Roeder, “Metabolism is the tie: The Bertalanffy-type cancer growth model as common denominator of various modelling approaches”, Biosystems, 167, 1–23, 2018. doi: 10.1016/j.biosystems.2018.03.004
44. M. Kühleitner, N. Brunner, W.G. Nowak, K. Renner-Martin, K. Scheicher, “Best fitting tumor growth models of the von Bertalanffy–Pütter type”, BMC Cancer, 19(1), 683, 2019. doi: 10.1186/s12885-019-5911-y
45. S.S. Hassan, H.M. Al-Saedi, “Comparative study of tumor growth based on single species models”, BIO Web of Conferences, 97, 00118, 2024. doi: 10.1051/bioconf/20249700118
46. L. Norton, “A Gompertzian model of human breast cancer growth”, Cancer Research, 48, 7067–7071, 1988.
47. C. Guiot, P.G. Degiorgis, P.P. Delsanto, P. Gabriele, T.S. Deisboeck, “Does tumor growth follow a ‘universal law’?”, Journal of Theoretical Biology, 225, 147–151, 2003. doi: 10.1016/S0022-5193(03)00221-2
48. R.G. Cornell (ed.), Statistical Methods for Cancer Studies, Statistics: A Series of Textbooks and Monographs, Vol. 51, CRC Press, 1984, 496 p.
49. F. Emmert-Streib, M. Dehmer (eds.), Statistical Diagnostics for Cancer: Analyzing High-Dimensional Data, Quantitative and Network Biology Series, 2013, 292 p.

50. H. Ishwaran, U.B. Kogalur, E.H. Blackstone, M.S. Lauer, “Random survival forests”, The Annals of Applied Statistics, 2(3), 841–860, 2008. doi: 10.1214/08-AOAS169
51. M. Kamran, J.Y. Abdullah, A.S. Ahmad Satmi, M. Genisa, A. Majeed, T. Nadeem, “Mathematical modeling and analysis of tumor growth models integrating treatment therapy”, Mathematics and Computers in Applications, 30, 119, 2025. doi: 10.3390/mca30060119
52. Y. Suter, M. Lê, G. Glauser, et al., “The LUMIERE dataset: Longitudinal glioblastoma MRI with expert RANO evaluation”, Scientific Data, 9(1), Article 768, 2022. doi: 10.1038/s41597-022-01881-7

Tiurdo Ivan *PhD of the Department of Applied Mathematics, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine;*

Sediuk Anastasiia *Student of the Faculty of Mathematics and Informatics, Department of Applied Mathematics, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine*

Kizilova Natalya *Doctor of Physical and Mathematics Sciences, Professor; professor of the Department of Applied Mathematics, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine;*

Mathematical modeling of tumor growth dynamics for personalized therapy selection.

Purpose of the work: to analyze current approaches to mathematical modeling of tumor growth and prediction of their dynamics using classical deterministic models and machine learning methods, as well as to determine the prospects for their use in modern mathematical oncology and personalized antitumor therapy.
Research methods: analysis and systematization of modern scientific publications on mathematical oncology; use of mathematical modeling methods for tumor growth (exponential, logistic, Gompertz and Bertalanffy models); statistical analysis of clinical data; application of machine learning methods for regression analysis and prediction of tumor growth dynamics based on longitudinal MRI data from the open LUMIERE dataset.
As a result of the study, a review and comparative analysis of classical mathematical models of tumor growth and their modifications used to describe the biological processes of proliferation and restriction of tumor tissue growth was performed. Preliminary processing and analysis of clinical and imaging data, including the volumes of various tumor components, was carried out. Individual tumor growth trajectories were modeled using regression models and ensemble machine learning methods, in particular Random Forest. It was shown that machine learning methods provide more stable and accurate predictions of complex tumor growth dynamics compared to classical models in the case of high data variability.
Conclusions: Combining classical mathematical models of tumor growth with modern machine learning methods is a promising direction for the development of mathematical oncology. This approach improves the accuracy of predicting individual tumor dynamics and creates a basis for developing personalized treatment strategies. The results obtained indicate the feasibility of further use of hybrid models in research on precision medicine and personalized antitumor therapy.
Keywords: *mathematical oncology, tumor growth models, Gompertz model, Bertalanffy model, machine learning, Random Forest, prediction, personalized therapy, precision medicine, clinical data.*

Надійшла у першій редакції 27.01.2026; рекомендована до друку 03.03.2026; опублікована 03.04.2026.
The first version has been received on 27.01.2026; was recommended for printing on 03.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-08

УДК (UDC) 519.8:004.72.052.4:004.8

Турчак Денис Сергійович *Аспірант кафедри прикладної математики Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, м. Харків, 61022*
e-mail: denys.turchak@karazin.ua
https://orcid.org/0009-0003-9144-8280

Руккас Кирило Маркович *Дтн , доцент Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, м. Харків, 61022*
e-mail: rukkas@karazin.ua;
https://orcid.org/0000-0002-7614-0793

Вплив архітектури GNN на робастність мережевих маршрутів у сценаріях одиничних відмов вузлів

У роботі розглянуто задачу інтелектуальної маршрутизації (пошуку шляху) в програмно-конфігурованих мережах (SDN) з використанням графових нейронних мереж. з метою підвищення ефективності використання мережевих ресурсів та адаптації до динамічних змін стану мережі (наприклад, завантаженості каналів або затримок). **Актуальність.** Сучасні програмно-конфігуровані мережі (SDN) стикаються зі зростанням обсягів трафіку та підвищеними вимогами до якості обслуговування (QoS). Традиційні алгоритми маршрутизації (наприклад, Дейкстри) є статичними та неефективними в умовах високої динаміки навантаження або раптових змін топології (відмов обладнання). Це призводить до перевантаження каналів, збільшення затримок та втрат пакетів. Використання методів машинного навчання, зокрема графових нейронних мереж (GNN) та навчання з підкріпленням (RL), відкриває нові можливості для створення адаптивних інтелектуальних агентів, здатних оптимізувати маршрутизацію в реальному часі, що робить це дослідження своєчасним та важливим для розвитку телекомунікаційних систем.
Мета. Метою дослідження є підвищення ефективності та надійності передачі даних у мережах SDN шляхом розробки та порівняльного аналізу інтелектуальних методів маршрутизації. Основний фокус зосереджено на дослідженні архітектур графових нейронних мереж (GCN, GAT) та алгоритму Q-Learning для забезпечення адаптивного керування трафіком в умовах змінного навантаження та відмов мережевих вузлів.
Методи дослідження. Методологічна основа роботи базується на комплексному застосуванні теорії графів для формалізації топології мережі, методів глибокого навчання для обробки ознак вузлів та алгоритмів навчання з підкріпленням для прийняття рішень щодо маршрутизації. Експериментальна верифікація запропонованих підходів здійснювалася шляхом емуляції програмно-конфігурованої мережі у середовищі Mininet під управлінням контролера Ryu. Програмна реалізація включала розробку моделей на основі згорткових мереж (GCN) та мереж з механізмом уваги (GAT) з використанням бібліотек глибокого навчання, а також імплементацію агента Deep Q-Learning. Оцінка ефективності алгоритмів проводилася за допомогою порівняльного аналізу ключових метрик якості обслуговування — пропускної здатності, середньої затримки та відсотка втрат пакетів — у сценаріях поступового зростання навантаження та аварійної зміни топології внаслідок відмови обладнання.
Результати. У ході дослідження встановлено, що інтеграція методів машинного навчання дозволяє суттєво покращити параметри передачі даних порівняно з класичним алгоритмом Дейкстри, особливо в умовах високого трафіку, де інтелектуальні агенти забезпечують меншу затримку та стабільність з'єднання. Критичний аналіз стійкості до відмов виявив суттєві розбіжності між досліджуваними архітектурами: модель GCN продемонструвала обмежену здатність до адаптації з показником невдалих спроб маршрутизації на рівні 30%, тоді як архітектура GAT показала кращу гнучкість, формуючи оптимальні шляхи у половині випадків. Найвищу ефективність підтвердив метод Q-Learning, який завдяки динамічній взаємодії із середовищем забезпечив побудову ідеальних маршрутів у 85% експериментів та мінімізував втрати пакетів до 5–7% навіть у критичних ситуаціях, що доводить перевагу підходів Reinforcement Learning над методами Supervised Learning у задачах адаптивного керування мережами.

Ключові слова: *Програмно-конфігуровані мережі (SDN), Графові нейронні мережі (GNN), Навчання з підкріпленням (Reinforcement Learning), Інтелектуальна маршрутизація, GCN, GAT, Q-Learning, Якість обслуговування (QoS), Відмовостійкість, Адаптивне керування трафіком.*

Як цитувати: Турчак Д. С Руккас. К. М., Вплив архітектури GNN на робастність мережевих маршрутів у сценаріях одиничних відмов вузлів. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.101-110. https://doi.org/10.26565/2304-6201-2026-69-08

How to quote: D. Turchak, K. Rukkas,“The impact of GNN architecture on the robustness of edge routes in scenarios of single node types”, *Bulletin of V. N. Karazin Kharkiv National University, series*

Mathematical modelling. Information technology. Automated control systems, vol. 69, pp. 101-110, 2026. https://doi.org/10.26565/2304-6201-2026-69-08 [in Ukrainian]

1 Вступ

Використання нейронних мереж для побудови маршрутів - це одна з найбільш динамічних галузей, у сферах програмно-конфігурованих мереж (SDN), логістики та автономної навігації. Традиційні алгоритми (як-от Дейкстра або A*) ефективні, але вони часто не встигають за динамічними змінами в реальному часі. Епоха 6G, масивних IoT-мереж та супутникового інтернету висуває нові вимоги до швидкодії та відмовостійкості маршрутизації. У 2025 році пошук "найкоротшого шляху" поступився місцем багатофакторній оптимізації, що враховує затримки, енергоспоживання та пріоритетність трафіку в реальному часі. Використання нейронних мереж дозволяє вийти за межі реактивного керування, впроваджуючи предиктивні моделі, які здатні розв'язувати конфлікти в мережі ще до їх виникнення.

Аналіз сучасних робіт показує, що наукова спільнота остаточно відійшла від використання нейромереж як «чорних скриньок»[3]. Сьогодні домінує підхід інтелектуальної маршрутизації, яка поєднує знання про топологію (GNN) та стратегічне прийняття рішень (DRL).

Сучасні підходи до побудови маршрутів базуються на інтеграції графових нейронних мереж (GNN) для аналізу топології та глибокого навчання з підкріпленням (DRL) для стратегічного прийняття рішень. Ключовим етапом стала еволюція моделей сімейства RouteNet (Fermi, Erlang[, які завдяки механізмам Message Passing та Attention забезпечують предиктивне моделювання QoS (затримки, джиттеру) з точністю пакетних симуляторів, але значно вищою швидкістю [1, 2, 6]. Паралельно розвиваються гібридні фреймворки, як-от Grace (2025) та багатоцільові моделі DQNR, де GNN формує ембеддінги стану мережі, а DRL-агенти динамічно оптимізують маршрути за критеріями пропускної здатності (приріст до 40%), енергоефективності та стабільності лінків [4, 5, 9]. Новітні розробки також залучають каузальний висновок та архітектури LLM-NAR для підвищення логічної прозорості та масштабованості інтелектуальних систем керування у складних 3D-мережах [7].

Більшість сучасних досліджень фокусується на максимізації пропускної здатності та мінімізації затримок, залишаючи аспект відмовостійкості на другому плані. Проте для критично важливих застосунків здатність алгоритму до самовідновлення є пріоритетною над його обчислювальною ефективністю. З огляду на це, постає необхідність оцінити, наскільки існуючі GNN- та DRL-алгоритми є фактично стійкими до відмов у системі, та якою мірою вони здатні самостійно знаходити обхідні маршрути при зміні топології.

2 Середовище та об'єкт дослідження

Експериментальне середовище побудоване на інтеграції емулятора Mininet та SDN-контролера Ryu, що функціонують у закритому циклі керування трафіком. Основна увага приділена механізмам передачі стану мережі від фізичного рівня до інтелектуального агента.

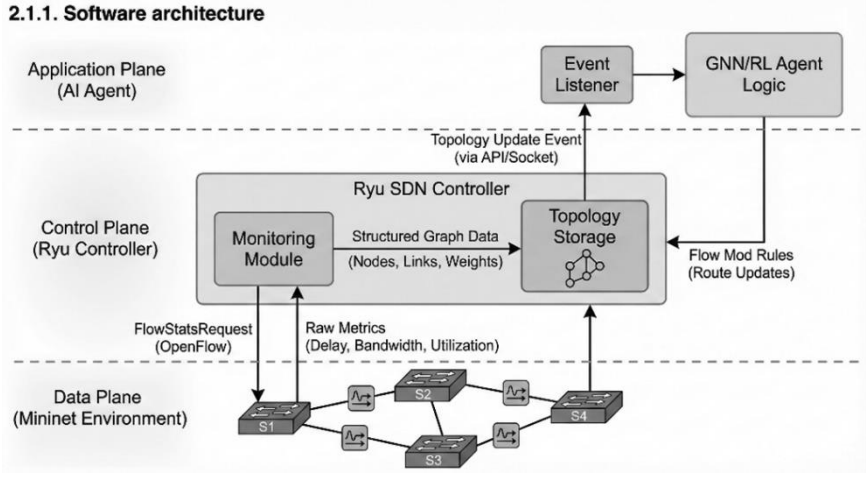


Рис 1. Архітектура експериментального середовища.
Fig. 1. Architecture of the experimental environment.

Контролер Ryu виконує роль центрального вузла збору метрик. Взаємодія реалізована через наступні модулі:

1. Модуль моніторингу: У реальному часі здійснює опитування комутаторів (FlowStatsRequest) для отримання значень затримки, доступної смуги пропускання та рівня використання лінків
2. Сховище топології: Спеціалізований буфер даних, що агрегує сирі метрики з площини даних та трансформує їх у структурований опис графа. Це дозволяє абстрагувати топологію від особливостей протоколу OpenFlow.
3. Event Listener: Програмний інтерфейс у площині застосунків, який реагує на зміни в Topology Storage. Щойно фіксується оновлення (наприклад, зміна завантаження лінка або відмова вузла), агент ініціює перерахунок маршрутів.

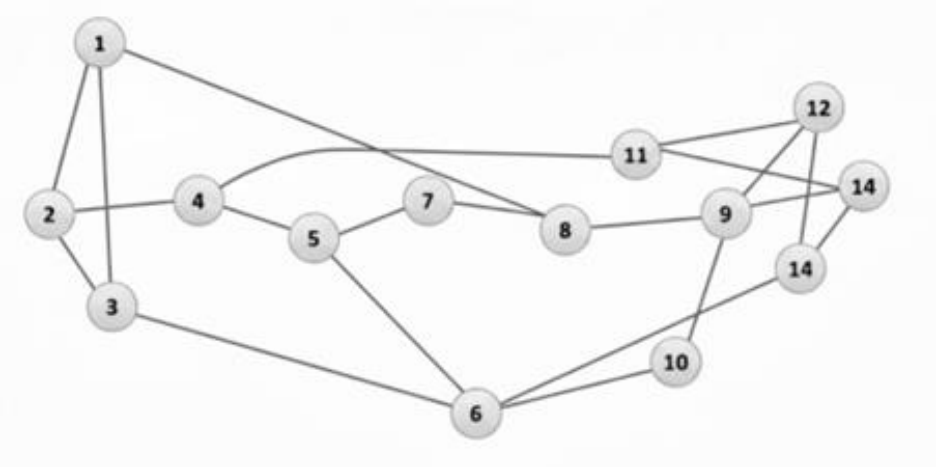


Рис 2. Мережа Національного Наукового Фонду
Fig 2. National Science Foundation Network

Для проведення експериментів було обрано топологію NSFNET (National Science Foundation Network), Цей вибір є класичним для задач оптимізації маршрутизації, оскільки структура мережі містить декілька замкнених контурів, що дозволяє інтелектуальним агентам обирати обхідні шляхи при відмові критичних вузлів. Мережа формалізується як неорієнтований граф $G_t = (V, E, W_t)$. Множина вузлів V представляє OpenFlow комутатори. Кожен вузол має вектор ознак – обсяг буфера та поточне завантаження CPU контролера. Множина ребер E представляє фізичні або віртуальні коанали зв'язку. Матриця ваг W_t представляє динамічні характеристики лінків, такі як затримка, джиттер та доступна смуга пропускання.. На відміну від топології IEEE 39-bus[8], яка використовується для специфічних кіберфізичних систем електромереж, обрана нами топологія NSFNET дозволяє тестувати моделі в умовах магістральної мережі передачі даних загального призначення.

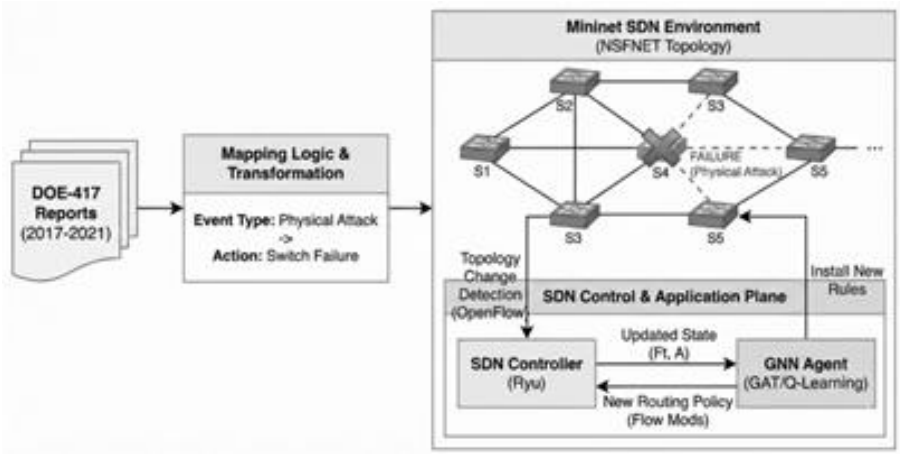


Рис 3. Процес симуляції трафіку на основі реальних звітів про відмови
Fig. 3. Traffic simulation process based on real failure reports

Процес генерації трафіку в середовищі Mininet базується на поєднанні стандартних інструментів вимірювання (iperf) та методів сокетного програмування для відтворення динамічних патернів навантаження. На відміну від традиційного використання синтетичних генераторів трафіку, у даній роботі моделювання навантаження базується на реальних історичних даних про мережеві збої та надзвичайні ситуації. Для відтворення динаміки трафіку використано звіти **Electric Emergency and Disturbance Report (DOE-417)** за період з 2017 по 2021 роки. Ці звіти містять часові мітки початку та завершення інцидентів, типи подій (вандалізм, системні збої, екстремальні погодні умови) та опис зон ураження. За допомогою сокетного програмування в Mininet реалізовано механізм «програвання» цих подій, де кожен тип аварії з репорту трансформується у відповідний сплеск трафіку або відмову вузла в топології NSFNET.

Події типу «системне перевантаження» або «затримка передачі» використовуються для моделювання стохастичного притоку пакетів (Event-Driven traffic). Події типу «фізична атака» або «диверсія» імітують повну відмову комутаторів (switch failure). Кожен сплеск черги на комутаторі, спричинений подією з DOE-417, перетворюється на функцію часу, яка вказує на наближення до критичного порогу заповнення буфера.

Такий підхід дозволяє верифікувати робастність запропонованих моделей (GAT, GCN) у сценаріях, які максимально наближені до реальних умов експлуатації магістральних мереж під час критичних збоїв.

Для навчання нейромережі використовуються три типи вхідних тензорів. Тензор трафіку D_t – матриця вимог «джерело-призначення», де кожен елемент d_{ij} вказує обсяг трафіку який необхідно передати між вузлами i та j. Тензор стану топології S_t включає інформацію про працездатність елементів. Вектор відмов F_t- бінарна маска $f \in \{0,1\}^{|E|+|V|}$, де 1 сигналізує про відмову конкретного вузла або лінка.

Завдання інтелектуальної системи полягає у знаходженні оптимальної стратегії маршрутизації π, яка відображає поточний стан мережі S_t у множину дій a_t(вибір шляхів), мінімізуючи сумарні втрати при дотриманні жорстких обмежень QoS.

Ми розглядаємо задачу багатоцільової оптимізації. Агент (нейромережа) повинен мінімізувати функцію вартості $J(\pi)$ на часовому горизонті T.

$$J(\pi) = \mathrm{E}_{\pi}[\sum_{t=0}^{T} \gamma^t C(S_t, a_t)] \tag{2.1}$$

Де миттєва вартість C визначається як зважена сума критичних метрик:

$$C = \omega_1 \text{Затримка}_{\text{загальна}} + \omega_2 * \text{Втрата пакетів} + \omega_3 * \frac{1}{\text{Надійність}} \quad \gamma \in [0,1] \tag{2.2}$$

Де γ – коефіцієнт дисконтування, а ω_i- вагові коефіцієнти, що визначають пріоритет (наприклад, для VoIP-трафіку ω_1 буде максимальним).

У формулі цільової функції (Cumulative Reward) γ використовується для зважування послідовності нагород:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} \ldots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{2.3}$$

Де R_t- негайна нагорода, $\gamma^k R_{t+k+1}$- нагорода, яку агент отримає через k кроків, помножена на коефіцієнт у ступені k. При побудові маршруту $P_{s,d}$ нейромережа повинна задовольняти наступні умови:

Обмеження пропускної здатності:

Для кожного ребра $e \in E$

$$\sum_{f \in \phi} x_e^f R_f \leq B_e \, (1 - f_e) \tag{2.4}$$

де x_e^f- індикатор використання лінка потоком f, R_f- швидкість потоку, B_e – смуга пропускання, f_e- стан відмови.

Умова цілісності потоку: Для кожного вузла $v \in V$:

$$\sum_{j:(v,j)\in E} f_{v,j} - \sum_{i:(i,v)\in E} f_{i,v} = \begin{cases} 1, \text{якщо } v = source \\ -1, v = destination \\ 0, \text{в іншому випадку} \end{cases} \quad (2.5)$$

Результатом роботи моделі є розподіл ймовірностей над множиною можливих наступних хопів або вибір повного шляху з набору K-найкоротших шляхів, що забезпечує максимальну стійкість до зафіксованого вектора відмов F_t

У межах цього завдання планується оцінити роботу кількох архітектур —GCN, GAT та GNN-DRL — з урахуванням їхніх принципів навчання, структури та способів взаємодії з контролером SDN. Для цього передбачається використання симуляційних середовищ (Mininet) які дозволяють змінювати стан топології в реальному часі, моделювати відмови та вимірювати ключові показники якості обслуговування (QoS), такі як середня затримка, джитер, пропускна здатність і швидкість відновлення маршрутизації.

3 Моделі

Для вирішення задачі регресії на вузлах графа було розглянуто та порівняно дві основні архітектури, що базуються на графових нейронних мережах (GNN): згорткову графову мережу (GCN) та графову мережу уваги (GAT). Усі моделі приймають стандартизований набір вхідних даних: матрицю ознак вузлів$X \in R^{N\times(F+N)}$ де N – кількість вузлів, F – кількість ознак) та зважену матрицю суміжності $A \in R^{N\times N}$. Задачею є прогнозування одного скалярного значення для кожного вузла, тобто $Y' \in R^{N\times 1}$.

3.1. Модель 1: Двошарова GCN (Baseline)

Рис 4. Архітектура GCN для вирішення задачі маршрутизації.
Fig. 4. GCN architecture for solving the routing problem

В якості базової моделі (baseline) використано стандартну двошарову архітектуру GCN. Модель складається з двох послідовних шарів GCNConv, кожен з яких використовує ReLU як функцію активації та генерує EMBEDDING_DIM ознак. Принцип агрегації: GCN виконує ізотропну (статичну) агрегацію. Інформація від сусідніх вузлів усереднюється з вагами, що жорстко визначені нормалізованою матрицею суміжності Â. Цей механізм є обчислювально ефективним, однак він не розрізняє важливість сусідів у залежності від їхніх ознак.

3.2. Модель 2: Тришарова GAT (Attention-based)

Рис 5. Архітектура GAT для вирішення задачі маршрутизації.
Fig. 5. GAT architecture for solving the routing problem.

Для врахування динамічної важливості сусідніх вузлів було реалізовано глибшу, тришарову модель на основі GATConv (Graph Attention Network), зокрема її покращеної версії GATv2. Шари 1 та 2: Використовують 4 голови уваги (attn_heads=4). Виходи голів конкатенуються, що призводить до розширення простору ознак до $\mathbb{R}^{N\times EMBEDDING\_DIM\times 4}$. Шар 3 (Фінальний): Використовує 8 голів уваги (attn_heads=8). Виходи голів усереднюються (concat_heads=False), повертаючи тензор до розмірності $\mathbb{R}^{N\times EMBEDDING\_DIM}$. Принцип агрегації: GAT виконує анізотропну (динамічну) агрегацію. Модель обчислює коефіцієнти уваги для кожної пари з'єднаних вузлів "на льоту", базуючись на їхніх поточних ознаках. Це дозволяє моделі навчитися призначати вищу вагу більш релевантним сусідам та ігнорувати менш важливі, що є ключовою перевагою для складних графів.

4 Тренування

У рамках проведеного дослідження стратегія навчання інтелектуальних агентів реалізована з використанням двох комплементарних підходів: навчання з підкріпленням на основі алгоритму Q-learning для динамічної адаптації у середовищі та навчання з учителем на основі попередньо підготовленого набору даних.

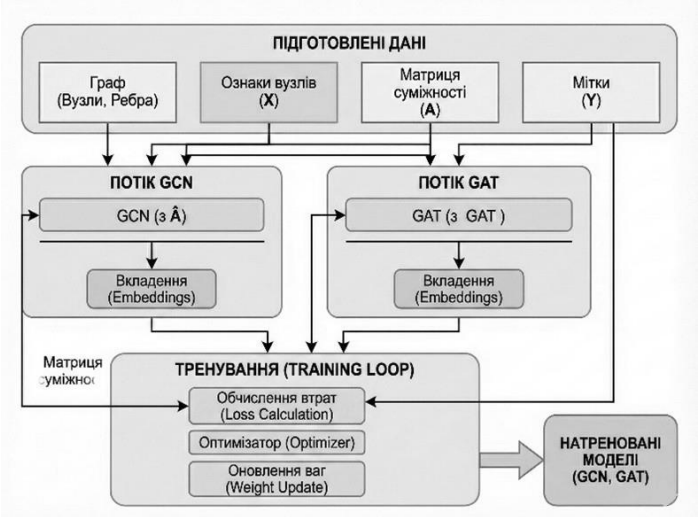


Рис 6. Алгоритм тренування моделей в навчанні з підкріпленням.
Fig. 6. Algorithm for training models in reinforcement learning.

4.1 Навчання з підкріпленням

Процес навчання ініціюється етапом підготовки даних, де вхідна інформація декомпозується на структурні елементи графа (вузли та ребра) та матричні представлення. Ключовими

компонентами тут виступають матриця ознак вузлів (X), що містить атрибутивну інформацію, та матриця суміжності (A), яка формалізує топологію зв'язків мережі. Для контрольованого навчання також використовуються цільові мітки (Y), що слугують еталоном (ground truth) для оцінки точності прогнозів.

4.2 Q-learning

Процес прийняття рішень агентом у графі реалізовано на основі алгоритму навчання з підкріпленням (Deep Q-Learning), де навігація моделюється як послідовність переходів між вузлами мережі. На кожному часовому кроці t поточний стан агента S_t характеризується набором локальних ознак X_t, які попередньо обробляються модулем графових нейронних мереж (GCN або GAT). Цей етап дозволяє трансформувати сирі дані у компактні векторні представлення (embeddings), що кодують структурну інформацію про оточення вузла.

Навчання системи відбувається через механізм зворотного зв'язку: після виконання дії агент отримує сигнал винагороди R_{t+1}, який кількісно визначає успішність переходу. Ця величина разом із максимальною оцінкою нового стану використовується блоком «Q-Learning Оновлення» для обчислення помилки передбачення та коригування вагових коефіцієнтів MLP. Такий ітеративний підхід дозволяє агенту поступово оптимізувати стратегію навігації, навчаючись обирати маршрути, що максимізують довгострокову винагороду.

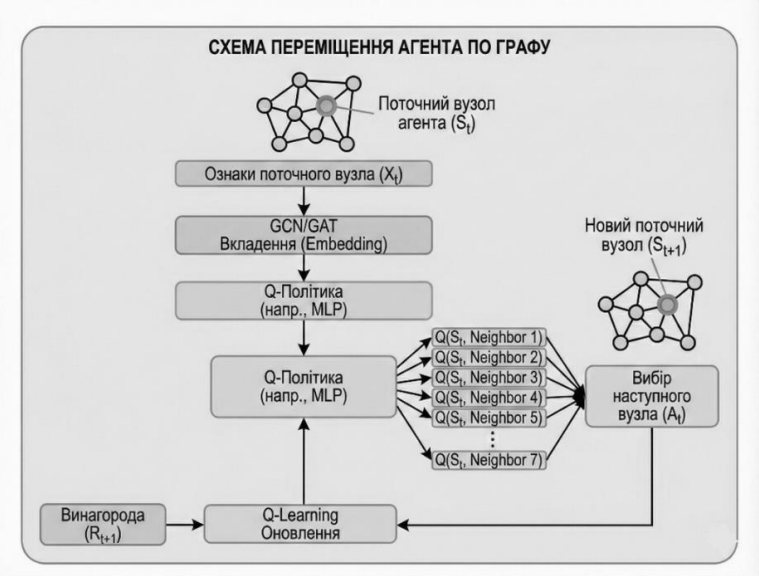


Рис 7. Алгоритм тренування моделі при Q-Learning підході.
Fig. 7. Model training algorithm using the Q-Learning approach.

5 Результати

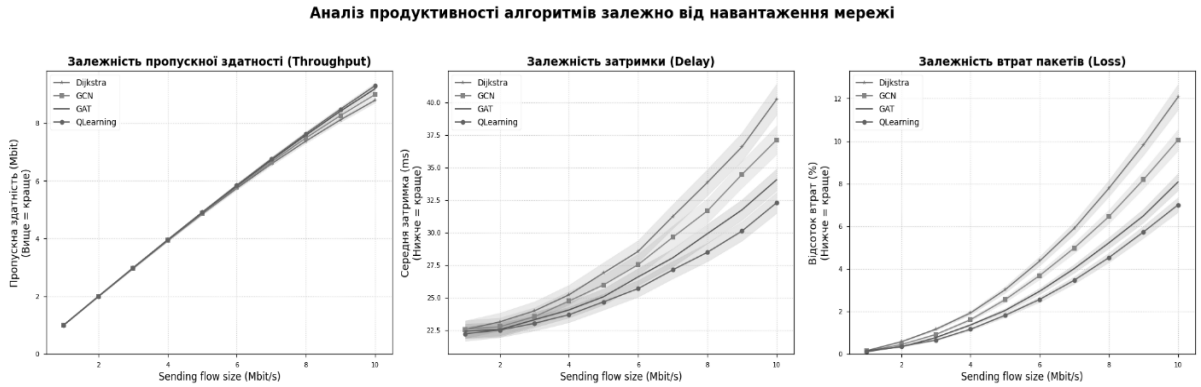


Рис 8. Порівняння ефективності моделей на тренованій топології.
Fig. 8. Comparison of the effectiveness of models on the trained topology.

Результати експерименту демонструють порівняльну ефективність методів маршрутизації (Дейкстра, GCN, GAT, Q-Learning) при зростанні навантаження мережі від 1 до 10 Мбіт/с. Хоча пропускна здатність (Throughput) залишається зіставною для всіх підходів, показники якості

обслуговування суттєво різняться при збільшенні інтенсивності потоку. Статичний алгоритм Дейкстри виявляє найшвидшу деградацію параметрів, досягаючи критичних значень затримки та втрат пакетів (понад 12%) у пікових режимах.

Серед інтелектуальних методів найкращу масштабованість та стійкість до перевантажень продемонстрував агент Q-Learning, який забезпечив мінімальну затримку та найнижчий відсоток втрат (близько 7%). Архітектури на базі GNN також перевершили класичний підхід, причому механізм GAT виявився ефективнішим за GCN. Отримані дані підтверджують перевагу динамічної адаптації маршрутів на основі навчання з підкріпленням над статичними та суто прогнозними методами.

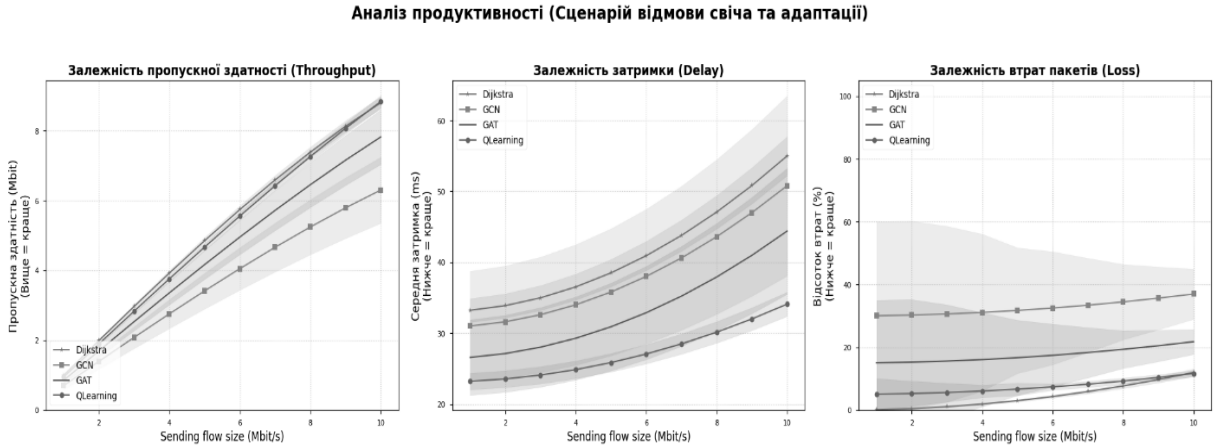


Рис 9. Порівняння ефективності моделей на тренованій топології при відмові вузлів.
Fig. 9. Comparison of the effectiveness of models on the trained topology in the event of node failure.

Експеримент із моделюванням відмови комутатора виявив критичні відмінності в адаптивності досліджуваних методів. На відміну від штатного режиму роботи, де показники були близькими, зміна топології спричинила значну деградацію продуктивності моделей на основі GNN. Зокрема, модель GCN продемонструвала найгірші результати: суттєве зниження пропускної здатності та стабільно високий рівень втрат пакетів (близько 30%) незалежно від навантаження, що свідчить про низьку здатність до генералізації при зміні структури графа. Архітектура GAT показала кращу адаптивність (втрати близько 15%), проте також не змогла ефективно перебудувати маршрути.

Рис 10. Співвідношення якості шляхів у сценарії відмови вузла.
Fig. 10. Path quality correlation in node failure scenario.

Найвищу відмовостійкість продемонстрував агент Q-Learning, який не лише зберіг пропускну здатність на рівні класичного алгоритму Дейкстри, але й забезпечив найкращі показники QoS. У той час як алгоритм Дейкстри через перевантаження резервних каналів показав найбільшу затримку, Q-Learning успішно знаходив оптимальні обхідні шляхи, утримуючи мінімальні

значення затримки та втрат пакетів навіть при максимальному навантаженні. Це підтверджує перевагу динамічної стратегії RL над статичними методами та класичними GNN у критичних сценаріях.

На діаграмі наведено порівняльний аналіз ефективності маршрутизації для трьох інтелектуальних агентів (GCN, GAT, Q-Learning) в умовах аварійної зміни топології мережі. Результати класифіковано за трьома категоріями: ідеальні маршрути, субоптимальні (неідеальні) шляхи та невдалі спроби встановлення з'єднання.

Отримані дані свідчать про обмежену здатність моделі GCN до адаптації: лише 20% побудованих маршрутів були оптимальними, тоді як частка критичних збоїв досягла 30%. Архітектура GAT продемонструвала кращу стійкість, підвищивши відсоток ідеальних шляхів до 50% та знизивши рівень відмов удвічі (до 15%). Найвищу ефективність виявив алгоритм Q-Learning, який у 85% випадків забезпечив побудову ідеального маршруту, мінімізувавши частку невдалих спроб до 5%. Це підтверджує перевагу динамічного навчання з підкріпленням над підходами supervised learning при роботі з непередбачуваними змінами структури мережі.

6 Висновки

У роботі проведено комплексне дослідження ефективності методів інтелектуальної маршрутизації у програмно-конфігурованих мережах (SDN), порівнюючи класичний алгоритм Дейкстри, моделі на основі графових нейронних мереж (GCN, GAT) та підхід навчання з підкріпленням (Q-Learning). Результати моделювання підтвердили, що інтеграція машинного навчання дозволяє суттєво покращити показники якості обслуговування (QoS) порівняно зі статичними алгоритмами, особливо в умовах високого навантаження мережі.

Експериментальні дані засвідчили, що у стабільному режимі роботи архітектури GAT та Q-Learning демонструють найкращу масштабованість, забезпечуючи низьку затримку та мінімізацію втрат пакетів при зростанні трафіку до 10 Мбіт/с. При цьому механізм уваги в GAT виявився ефективнішим за спектральну згортку GCN, дозволяючи точніше враховувати вагу зв'язків між вузлами. Однак найбільш показовими стали результати тестування в аварійних сценаріях зі зміною топології мережі (відмова комутатора).

Ключовим результатом дослідження є доведення переваги динамічного підходу Q-Learning над методами навчання з учителем у нестабільних середовищах. У той час як моделі GNN зіткнулися з проблемою індуктивного узагальнення, продемонструвавши високий рівень невдалих маршрутів (до 30% для GCN), агент Q-Learning успішно адаптувався до нових умов. Завдяки механізму безперервної взаємодії з середовищем, він забезпечив побудову ідеальних шляхів у 85% випадків, утримуючи параметри продуктивності на рівні штатного режиму. Перспективи подальших досліджень спрямовані на розробку гібридних архітектур, які дозволять об’єднати можливості GNN щодо аналізу просторових ознак топології з адаптивною гнучкістю методів навчання з підкріпленням. Окремим вектором розвитку є оптимізація швидкості збіжності алгоритмів та їх масштабування для великих магістральних мереж зі складними ієрархічними зв’язками.

СПИСОК ЛІТЕРАТУРИ

1. Rusek K., et al. RouteNet: Leveraging Graph Neural Networks for network modeling and optimization. IEEE Journal on Selected Areas in Communications. 2020. Vol. 38, no. 10. P. 2260–2270. DOI: 10.1109/JSAC.2020.3000405.
2. Ferriol-Galmés M., et al. RouteNet-Erlang: A Graph Neural Network for Network Performance Evaluation. IEEE INFOCOM 2022 - IEEE Conference on Computer Communications. 2022. P. 2018–2027. DOI: 10.1109/INFOCOM48030.2022.9796677.
3. Mammeri A. A Survey on Machine Learning Techniques for Routing Optimization in SDN. IEEE Access. 2021. Vol. 9. P. 104523–104544. DOI: 10.1109/ACCESS.2021.3098763.
4. Stampa G., et al. A Deep Reinforcement Learning Approach for Software-Defined Networking Routing Optimization. IEEE International Conference on Communications (ICC). 2019. P. 1–6. DOI: 10.1109/ICC.2019.8761479.
5. Zheng S., Huang H., GNN et al. Research on Generalized Intelligent Routing Technology Based on Deep Reinforcement Learning. Electronics. 2022. Vol. 11, no. 3. P. 343. DOI: 10.3390/electronics11030343.

6. Ferriol-Galmés M., Barlet-Ros P., Cabellos-Aparicio A. RouteNet-Fermi: Network Modeling with Graph Neural Networks. IEEE/ACM Transactions on Networking. 2024. Vol. 32, no. 3. P. 1200–1215. DOI: 10.1109/TNET.2024.3392336.
7. Wu W., et al. Neural Algorithmic Reasoners informed Large Language Model for Multi-Agent Path Finding (LLM-NAR). arXiv preprint arXiv:2508.17971. 2025. URL: https://arxiv.org/abs/2508.17971 (дата звернення: 28.01.2026).
8. Islam M., et al. Results analysis on the cyber layer of the IEEE 39-bus test system under both normal and failure condition. IEEE Access. 2024. Vol. 12. P. 106234–106245. DOI: 10.1109/ACCESS.2024.10623446.
9. Grace: Toward Routing in Dynamic Network Environments With Graph Embedding. IEEE Transactions on Networking. 2025. DOI: 10.1109/TNET.2025.11078667.

Turchak Denys Serhiyovich *Postgraduate student of the Department of Applied Mathematics*
V. N. Karazin Kharkiv National University, Svobody Square, 4, Kharkiv, 61022
e-mail: denys.turchak@karazin.ua
https://orcid.org/0009-0003-9144-8280

Rukkas Kyrylo Markovych *Doctor of Technical Sciences, Associate Professor*
V. N. Karazin Kharkiv National University, Svobody Square, 4, Kharkiv, 61022
e-mail: rukkas@karazin.ua;
https://orcid.org/0000-0002-7614-0793

The impact of GNN architecture on the robustness of edge routes in scenarios of single node types

The paper considers the problem of intelligent routing (path finding) in software-defined networks (SDN) using graph neural networks in order to increase the efficiency of network resource use and adapt to dynamic changes in the network state (for example, channel congestion or delays).

Topicality. Modern software-defined networks (SDN) are faced with increasing traffic volumes and increased quality of service (QoS) requirements. Traditional routing algorithms (e.g. Dijkstra) are static and inefficient in conditions of high load dynamics or sudden topology changes (hardware failures). This leads to channel congestion, increased latency, and packet loss. The use of machine learning methods, in particular graph neural networks (GNN) and reinforcement learning (RL), opens up new opportunities for creating adaptive intelligent agents capable of optimizing routing in real time, which makes this research timely and important for the development of telecommunication systems.

Goal. The aim of the research is to improve the efficiency and reliability of data transmission in SDN networks by developing and comparative analysis of intelligent routing methods. The main focus is on the study of graph neural network architectures (GCN, GAT) and the Q-Learning algorithm to provide adaptive traffic management under conditions of variable load and network node failures.

Research methods. The methodological basis of the work is based on the integrated application of graph theory to formalize the network topology, deep learning methods to process node features, and reinforcement learning algorithms to make routing decisions. Experimental verification of the proposed approaches was carried out by emulating a software-configured network in the Mininet environment under the control of the Ryu controller. The software implementation included the development of models based on convolutional networks (GCN) and attention networks (GAT) using deep learning libraries, as well as the implementation of the Deep Q-Learning agent. The effectiveness of the algorithms was assessed by comparative analysis of key quality of service metrics - throughput, average latency, and packet loss percentage - in scenarios of gradual load growth and emergency topology change due to equipment failure.

The results. The study found that the integration of machine learning methods allows for significant improvements in data transmission parameters compared to the classic Dijkstra algorithm, especially in high traffic conditions, where intelligent agents provide lower latency and connection stability. Critical analysis of fault tolerance revealed significant differences between the studied architectures: the GCN model demonstrated limited adaptability with a routing failure rate of 30%, while the GAT architecture showed better flexibility, generating optimal paths in half of the cases. The highest efficiency was confirmed by the Q-Learning method, which, thanks to dynamic interaction with the environment, ensured the construction of ideal routes in 85% of experiments and minimized packet loss to 5–7% even in critical situations, which proves the superiority of Reinforcement Learning approaches over Supervised Learning methods in adaptive network control tasks.

Key words: Software-defined networks (SDN), Graph neural networks (GNN), Reinforcement learning, Intelligent routing, GCN, GAT, Q-Learning, Quality of service (QoS), Fault tolerance, Adaptive traffic management.

Надійшла у першій редакції 27.01.2026; рекомендована до друку 03.03.2026; опублікована 03.04.2026.
The first version has been received on 27.01.2026; was recommended for printing on 03.03.2026; published 03.04.2026.

DOI: https://doi.org/10.26565/2304-6201-2026-69-09

УДК (UDC) 004.056.5

Чепель Данило Олександрович
аспірант
кафедри кібербезпеки інформаційних систем, мереж і технологій, Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022;
e-mail: dan4epel@gmail.com
https://orcid.org/0009-0009-7449-8095

Малахов Сергій Віталійович
кандидат технічних наук, доцент
кафедри кібербезпеки інформаційних систем, мереж і технологій, Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022;
e-mail: malakhov@karazin.ua
https://orcid.org/0000-0001-8826-1616

Гончаров Микита Олександрович
аспірант
кафедри кібербезпеки інформаційних систем, мереж і технологій, Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022;
e-mail: m.honcharov@student.karazin.ua
https://orcid.org/0000-0002-9790-7260

Застосування парадигми прецедентного аналізу для цілей мультибазового хмарного моніторингу DNS-трафіку

Актуальність. Зростання складності DNS-інфраструктури та підвищення рівня загроз у мережевому середовищі зумовлюють необхідність розроблення інтелектуальних засобів моніторингу DNS-трафіку, здатних забезпечувати прозоре, адаптивне та обґрунтоване виявлення поведінкових аномалій. Особливої актуальності набуває впровадження підходів, що підвищують простежуваність логіки прийняття рішень системами штучного інтелекту.
Мета. Метою роботи є експериментальне дослідження прототипу програмного засобу для моніторингу поточного стану DNS-трафіку з широкою імплементацією можливостей ШІ, в основу логіки якого покладено концепцію прецедентного аналізу (CBR) поведінкових аномалій DNS-трафіку.
Методи дослідження. У роботі використано методи імітаційного моделювання, мультибазові вимірювання часу обробки DNS-запитів із застосуванням системи розподілених хмарних датчиків-тестерів, а також алгоритми прецедентного аналізу для інтелектуальної постобробки даних. Прототип реалізовано у вигляді Python-клієнта, інтегрованого з Gemini API, що функціонує на основі набору даних, сформованого за результатами попередніх досліджень [1-2]. У процесі роботи система автономно модифікує реєстр аномалій шляхом додавання нових прецедентів на основі результатів аналітичної обробки.
Результати. Отримані результати демонструють, що розглянутий підхід для моніторингу DNS-трафіку забезпечує виявлення як уже відомих аномалій, так і локалізацію ще невідомих колізій. Підтверджено перспективність застосування прецедентного підходу для покращення оперативності корегувань параметрів діючої зони політики реагування (RPZ) [3] та підвищення рівня поінформованості персоналу з питань безпеки DNS-трафіку. Водночас експерименти виявили ефект т. зв. «кластеризації», що може призводити до хибнопозитивних результатів оцінки подій та, як наслідок, суперечливих трактувань отриманих відомостей щодо спостережуваних мережевих подій.
Висновки. Перегляд діючих обмежень та завдань аналізу для модулів ШІ і подальше моделювання підтвердили, що внесені зміни суттєвим чином зменшили виявлений ефект «кластеризації» та підвищили надійність інтерпретацій аномалій, які спостерігаються за визначеною системою непрямих (опосередкованих) ознак. Отримані результати підтверджують доцільність подальшого розвитку підходу прецедентного аналізу в системах інтелектуального моніторингу DNS-трафіку.

Ключові слова: *інформаційна безпека, штучний інтелект, фільтрація трафіку, DNS, RPZ, CBR, мережеві аномалії, хмарні обчислення, розподілена мережа, протоколи DNS.*

Як цитувати: Чепель Д. О., Малахов С. В., Гончаров М. О. Застосування парадигми прецедентного аналізу для цілеймультибазового хмарного моніторингу DNS-трафіку. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2026. вип. 69. С.111-121. https://doi.org/10.26565/2304-6201-2026-69-09
How to quote: D. Chepel, S. Malakhov and M. Honcharov, "Application of a preceden tanalysis paradigm for the purposes of multibase cloud monitoring of DNS traffic", *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 69, pp. 111-121, 2026. https://doi.org/10.26565/2304-6201-2026-69-09 [in Ukrainian]

1. Вступ

Науково-технічний прогрес у галузі інформаційних технологій сприяє безперервний розробці та впровадженню нових інформаційно-комунікаційних систем (ІКС) і технологій, але водночас, породжує й нові загрози безпеки, зокрема ті, що базуються на використанні DNS-орієнтованих векторів кібератак, такі як атаки DNS-ампліфікації, отруєння кешу резолверів та інші. За таких умов традиційні підходи до аналізу властивостей DNS-трафіку, виявляють досить серйозні функціональні обмеження. До таких умовно «сірих зон» можна віднести: – моніторинг зашифрованого чи обфускованого DNS трафіку; – спорадична динамічна зміна доменів (з боку зловмисників); – нерівномірність мережевих потоків; – робота в умовах впливу атак типу DNS Amplification тощо. Представлена робота є логічним продовженням цілої низки досліджень [1-3], спрямованих на вдосконалення можливостей (*перш за все, швидкість та інформативність*) хмарної мультибазової системи моніторингу поточного стану DNS-трафіку в заданих мережевих локаціях (доменних зонах). Основна увага зосереджена, насамперед, на оцінці застосовності парадигми прецедентного аналізу структури й властивостей поточного DNS трафіку з широким залученням можливостей ШІ. Попередньо, було зроблено припущення, що використання цього підходу може самим суттєвим чином підвищити обґрунтованість і оперативність коригування поточних параметрів зон політики реагування (*Response Policy Zones, RPZ*) та забезпечити своєчасне виявлення аномалій DNS-трафіку, що пов'язані з першими проявами загроз безпеки [1–4], котрі експлуатують вектор DNS. Під терміном «обґрунтованість», слід розуміти покращення ступеню поінформованості персоналу з питань ІБ про актуальний стан мережевих інцидентів, що експлуатують вектор DNS атак, через комплексність результатів аналізу даних моніторингу, в заданих мережевих сегментах (локаціях). Як вказано у роботах [1-2], цей процес відбувається, перш за все, шляхом глибокої інтеграції системи розгалужених програмних датчиків-сенсорів (див. Рис.1 в роботі [1]) й можливостей технології ШІ.

Основна ідея парадигми прецедентного аналізу полягає у формуванні і підтриманні в актуальному стані масиву формалізованих даних (*т.з. поведінкових шаблонів*) про типову активність DNS трафіку, котрі використовуються для інтерпретації поточних станів DNS трафіку у відповідності до інформації наявної бази знань (прецедентів). Впровадження такого механізму дій потребує перегляду структури раніше запропонованого тестового алгоритму [1-2], а також модифікації модулів збору та попередньої обробки даних. Реалізація вказаних змін забезпечує потрібні умови для коректного формування, зберігання та подальшого використання отриманих відомостей про прецеденти (колізії, аномалії, інциденти тощо). Проведене комп'ютерне моделювання підтвердило доцільність інтеграції механізмів прецедентного аналізу до структури хмарної системи моніторингу DNS трафіку та підтвердило її здатність фіксувати та обробляти широкий спектр поведінкових аномалій DNS-трафіку.

Мета роботи полягає в розгляді й узагальнені результатів тестового моделювання оновленого механізму прецедентного аналізу поточних даних DNS-трафіку, з застосуванням системи хмарних мультибазових вимірів. Оновлена концепція моніторингу DNS трафіку (доопрацьований алгоритм + нові елементи): - забезпечує комплексний характер відомостей, стосовно поточних мережевих подій (*в частині DNS трафіку*); - зменшує час реагування на інциденти безпеки, які експлуатують вектор DNS атак; - покращує валідність результатів аналізу поведінкових аномалій DNS-трафіку у визначених сегментах глобальної мережі.

2. Аналіз останніх досліджень і публікацій

Моніторинг DNS-трафіку є важливою складовою в загальної системі забезпечення безпеки сучасних ІКС. Своєчасне та точне виявлення аномалій DNS-трафіку дає змогу зменшити потенційну шкоду від загроз, що базуються на експлуатації DNS векторів атак, а також підвищити загальну адекватність адміністрування діючими RPZ [1-3]. Особливості питань аналізу каналів розвідки загроз (*Threat Intelligence Feeds*), механізмів корегування *RPZ* та методів протидії ботнет активності й шифрування DNS-трафіку, коротко розглянуті в [1]. Спираючись на результати аналізу останніх тенденцій, які пов'язані з застосуванням ШІ для цілей моніторингу і фільтрації DNS-трафіку, слід виділити кілька напрямів досліджень, що є релевантними для умов та завдань цієї роботи [5-12]:

2.1 Штучний інтелект в проблематиці аналізу даних

Інтелектуальні системи дають змогу автоматизовано отримувати структуровані знання з великих за обсягами та різнорідних джерел даних, перевершуючи традиційні - статистичні підходи за критеріями масштабованості й адаптивністю цих процесів. Помітною рисою сучасних профільних публікацій є, тренд на інтеграцію методів машинного навчання (ML), логіко-орієнтованих підходів та оптимізаційних методів/способів, як ключових елементів сучасних технологій аналізу даних (*в широкому сенсі цієї проблематики*). В цьому контексті, глибинне навчання визначається, як домінуюча аналітична парадигма. У літературі також, відзначаються певні успіхи в детектуванні й розпізнаванні зображень, стеганографії, обробці сигналів та виявленні мережевих аномалій, досягнуті нейронними мережами [6]. Водночас, попри високу точність прогнозування, глибинне навчання пов’язане з низкою обмежень, зокрема: – значними вимогами до обсягів даних, високою обчислювальною вартістю та низьким рівнем «прозорості» (*тобто, очевидності існуючих та/чи врахованих взаємозв'язків*). Це стимулює розвиток підходів, що поєднують доменно-орієнтовані знання, символічні методи та механізми переносу навчання з метою підвищення їх прикладної застосовності і прозорості систем ШІ. Крім того підкреслюється необхідність перевірки (валідації) відомостей, що згенеровано с залученням ШІ [6]. Загалом результати сучасних досліджень свідчать, що ШІ не лише підвищує ефективність та точність результатів аналізу різнорідних даних, але й трансформує самі підходи до механізмів реалізації та змісту процесів аналітики. В першу чергу це стосується зміщення уваги в бік синтезу інтегрованих інтелектуальних систем, котрі здатні оперативно обробляти складні мультимодальні дані в реальному масштабі часу. Попри швидкий прогрес, залишаються актуальними виклики, пов’язані з прозорістю моделей, необхідністю перевірки результатів та обчислювальними витратами [5-7].

2.2 Штучний інтелект в реаліях умов аналізу DNS-трафіку.

Сучасні фахівці з ІБ відзначають посилення тенденції на використання ШІ для завдань аналізу DNS-трафіку. Основною причиною є зниження ефективності традиційних сигнатурних (реактивних) підходів, які є малорезультативними в умовах інтенсивної обфускації каналів, швидкої зміни доменів та впровадження складних стратегій ухилення від виявлення. Дослідження показують, що моделі машинного навчання здатні виявляти аномалії мережевої поведінки навіть тоді, коли важливі інформативні ознаки приховані технологіями шифрування або навмисно маскуються [13-14]. Наголошується, що методи на основі ШІ забезпечують не лише вищу точність класифікації, але й здатність адаптуватися до нових умов та типів атак. Водночас проблема прозорості загальної логіки дій з боку ШІ, залишається вкрай актуальною. Вочевидь, що інтеграція функцій ШІ у реальні системи ІБ є неможливою без впровадження механізмів формалізованих інтерпретацій (пояснювання), стосовно змісту й логіки штучно синтезованих відомостей. Такий порядок валідації дій ШІ, дозволяє персоналу з безпеки зрозуміти, чому той чи інший запит, потік або процес було кваліфіковано, як підозрілий/аномальний. Ця потреба безпосередньо пов’язана з поняттям «довіри», можливістю аудиту прийнятих рішень і практичною придатністю таких систем у реальних умовах. Інакше кажучи, автоматизація процесів прийняття рішень з боку ШІ має гармонійно поєднуватися з процесом їх контролю. Це особливо важливо, в рамках потенційно неминучого зіткнення можливостей ШІ на боці протиборчих сторін («атака - захист»), де логіка дій та наслідки такої «взаємодії», виходять за рамки традиційних морально-етичних норм й впливу антропогенних

(фізіологічних) обмежень (наприклад, за кількість одночасно спостережуваних подій та/або інтенсивності їх появи та/чи генези тощо). Авторами ряду досліджень зазначається, що сучасні архітектури з залученням елементів ШІ, можуть працювати в режимі реального часу та залишатися стійкими до спроб ухилення з боку зловмисників. Це відкриває хороші перспективи, з точки зору розширення можливостей у сфері моніторингу DNS трафіку [8-9].

2.3 Прецедентний аналіз (Case-Based Reasoning)

Підхід «*Case-Based Reasoning»* (далі - *CBR*) набуває все більшої уваги, як основа для створення більш прозорих, адаптивних і зрозумілих для користувачів систем ШІ. Оскільки сучасні моделі ШІ дедалі все частіше демонструють поведінку у дусі умовної «чорної скриньки»,то саме концепція *CBR* розглядається як адекватний запобіжний механізм. Причина очевидна – така концепція дій апелює на досвід вже відомих випадків (прецедентів), тобто проеціює й масштабує логіку причинно-наслідкових аналогій (в даному контексті - штучного мислення). У дослідженні [10] підкреслюється, що базовий цикл *CBR: «пошук – повторне використання – коригування – збереження»,* забезпечує інтерпретовану структуру в якій результуючі рішення ШІ, ґрунтуються на минулому досвіді, а не на непрозорих статистичних залежностях. Це робить *CBR* особливо привабливим у сферах, де критично важливими є такі категорії й властивості, як: - довіра, переконливість (ґрунтовність дій) та можливість зворотного аудиту. Є думка [10-12], що прецедентні тлумачення підвищують рівень довіри користувачів і покращують розуміння рішень штучної інтелектуальної системи, особливо в експертних галузях, таких як охорона здоров'я чи оцінка поточного стану ІБ сучасних ІКС тощо.

Іншим перспективним напрямом досліджень є інтеграція *CBR* із сучасними моделями ШІ, зокрема з *великими мовними моделями* (*LLM*) [10,12]. В цьому разі *CBR* допомагає функціональним «агентам» на основі *LLM,* зменшувати кількість хибних трактувань, надійніше виконувати доменно-орієнтовані завдання та надавати обґрунтування своїх рішень/дій. Загалом, в якості проміжного висновку, можна стверджувати, що впровадження *CBR,* забезпечує як методологічні, так і концептуальні переваги для ШІ систем. Це підвищує прозорість, послідовність і адаптивність штучно синтезованих рішень, та робить поведінкову логіку систем ШІ, більш узгодженою з логікою мислення людини, наприклад: - психологічну схильність людей шукати підтвердження своїм вже ухваленим рішенням (що, з технічної точки зору, дуже співпадає з загальною концепцією *CBR*).

2. Основна частина

Враховуючи специфіку питань залучення можливостей ШІ до вирішення завдань різноманітних аналітичних систем, слід звернути увагу на той факт (підтверджений отриманими результатами моделювання), що логіка систем ШІ не є чимось унікальною і беззастережно аксіоматичною. У цьому контексті слід мати на увазі принцип (схему дій), що експлуатують кібершахраї при реалізації одного з різновидів атак соціального інжинірингу. Головне у такій схемі - це експлуатація впевненості жертви атаки, у своїй раціональності: - в частині виконуваних дій та/чи думок (логіки рішень). - Така особливість повною мірою збігається з прагненням системи ШІ, самооптимізуватися в процесі вирішення покладених на неї завдань, ігноруючи загальний контекст і взаємозв'язок подій/процесів, що оцінюються. Звідси, більшою мірою, і виникає природа помилково-позитивних спрацьовувань. Для нівелювання зазначених наслідків доводиться обмежувати подібне прагнення ШІ до самооптимізації, шляхом впровадження додаткових інструкцій та прямих заборон. Запорукою належного виконання подібних обмежувально-керуючих дій є можливість реалізації (підтримки) інверсного аудиту логіки прийнятих рішень з боку ШІ.

У межах проведеного циклу моделювань досліджувалися особливості використання парадигми *CBR* для виконання завдань моніторингу поточного стану DNS трафіку із залученням можливостей ШІ. В якості умовного індикатора успішності подібної інтеграції, виступала задача автоматичного доповнення й модифікації таблиці прецедентів про аномалії DNS трафіку. Набір даних, використаний в експерименті, було отримано з попередньої роботи [1]. Такий підхід забезпечив безперервність умов вимірювань й порівнянність результатів. Тестовий програмний стенд реалізований за допомогою *Python*-клієнта, який виконує запити через *Gemini API*. Таким чином, поточний реліз моделюючого алгоритму, в якості вихідних даних використовував заздалегідь підготовлені відомості початкової (стартової) таблиці прецедентів. В загальному

випадку, відповідна таблиця/реєстр містить сукупність базових випадків та відповідні їм інтерпретації. Фрагмент початкової таблиці прецедентів, наведено в Таб. 1.

Після отримання початкового набору прецедентів, дослідна система автоматично розширює таблицю прецедентів, додаючи нові відомості з коментарями (тлумаченнями явищ й процесів), стосовно причин їх додавання. Така логіка дій дозволяє використовувати раніш отримані знання під час інтерпретації даних нових спостережень (мультибазових вимірів [1]). Це підвищує точність, прозорість (*послідовність й взаємопов'язаність подій*) та відтворюваність результатів аналітики спостережуваного процесу, в даному разі – аномалій DNS трафіку [3].

Аналіз відомостей даних реєстру прецедентів, котрі були згенеровано дослідною системою ШІ (фрагмент реєстру див. в Табл.2), дозволяють констатувати наступне:

1 - діюча модель дій (*логіка ШІ*) поряд з виявленням аномалій, які вже є в реєстрі, фіксує і раніш невідомі випадки, доповнюючи результуючу таблицю новими записами. Наприклад, відсутність відповідей або помилки виконання тестових DNS-запитів [1,3]. Це свідчить про те, що експериментальна ШІ система підтримує не лише сигнатурне виявлення відомих прецедентів, але здатна фіксувати й нові поведінкові аномалії і колізії DNS трафіку;

Таблиця 1. Фрагмент стартової таблиці прецедентів (аномалій)
Table 1. Fragment of the initial table of precedents (anomalies)

Location	Server Name	Test Domain	Plain query time (ms)	DoH query time (ms)	DoT query time (ms)	*Comment*
JAPAN	Google	**nic.ar**	931	241	-	*PQ latency spike*
FINLAND	OpenDNS	**gov.za**	1318	20	214	*PQ latency spike*
FRANCE	Quad9-Reserve	**bbc.co.uk**	444	801	1322	*PQ, DTQ and DHQ latency spike*
ISRAEL	OpenDNS	**paris.fr**	986	1301	1043	*PQ, DTQ and DHQ latency spike*
ISRAEL	Quad9	**bbc.co.uk**	943	362	993	*PQ, DTQ and DHQ latency spike*

***Прим**: - в таблицях 1-2, "PQ"," DTQ" та "DHQ" означають час незашифрованих, DoT та DoH запитів відповідно.*

2 - використання прецедентної парадигми обробки даних, сприяє зменшенню кількості хибних позитивних спрацьовувань. В даному разі аналітик (зворотний аудит) може позначати некоректні чи нерелевантні відомості реєстру, поступово вдосконалюючи (корегуючи) механізми прийняття рішень системою ШІ.

3 - під час тестування першої версії дослідної системи було виявлено явище так званого «***гіперфокусування уваги***» моделі. Це особливо помітно у випадках, коли спостерігалися кластери аномалій одного типу (приклад див. нижче на Рис.1).

Зокрема, така реакція системи фіксувалась у множині записів, що відповідають різним географічним локаціям, із яких було проведено вимірювання та/або DNS-серверам.

Так, поява екстремального значення затримки для одного з DNS протоколів [3-4] у послідовних записах, часто призводила до того, що дослідна система класифікувала всі такі записи, як «*піки затримки для одного протоколу*». При цьому, належним чином не враховувалась динаміка інших параметрів затримки в межах тих самих записів. В інших випадках, система позначала події типу «*Відсутність відповіді*» як аномалії, навіть якщо було достеменно відомо, що запитуваний DNS сервер не підтримує відповідний протокол. Такі реакції спостерігалися в тих випадках, коли ці записи розміщувалися поруч із записами, які дійсно містили збої виконання DNS запитів.

Отримані результати свідчать про те, що *локальна контекстна схожість записів* та *уявна безперервність шаблонів* можуть домінувати над формальними правилами загальної логіки процесу, дозволяючи ШІ формувати власні інтерпретації для «сусідніх» випадків (записів).

Відповідна колізія логіки роботи ШІ фіксувалась, навіть якщо таке узагальнення суперечить явним інструкціям та/або призводить до ігнорування інших важливих атрибутів даних (рис. 1) .

Location	Server Name	Test Domain	Plain	DoH	DoT	*Comment*
ISRAEL	ControlD-Reserve	**bbc.co.uk**	293	-	792	*No support for DHQ protocol*
ISRAEL	ControlD-Reserve	**sina.com.cn**	291	-	775	*No support for DHQ protocol*
ISRAEL	ControlD-Reserve	**nic.ar**	291	-	760	*No support for DHQ protocol*
ISRAEL	ControlD-Reserve	**gov.za**	293	-	785	*No support for DHQ protocol*
ISRAEL	ControlD-Reserve	**terra.com.br**	281	-	744	*No support for DHQ protocol*

Ignoring constraints that define supported protocols

Location	Server Name	Test Domain	Plain	DoH	DoT	*Comment*
FRANCE	Quad9-Reserve	**bbc.co.uk**	296	0	1335	*DHQ is 0*
FRANCE	Quad9-Reserve	**sina.com.cn**	758	0	1279	*DHQ is 0*
FRANCE	Quad9-Reserve	**nic.ar**	501	0	1324	*DHQ is 0*
FRANCE	Quad9-Reserve	**gov.za**	621	0	1282	*DHQ is 0*
FRANCE	Quad9-Reserve	**terra.com.br**	398	0	1330	*DHQ is 0*

Ignoring high plain and DoT latency

Рис.1 Фрагмент записів реєстру з «гіперфокусуванням уваги» ШІ
Fig.1 Fragment of registry entries with AI "hyperfocus of attention"

З метою підвищення репрезентативності даних, тестові вимірювання, з яких був зібраний датасет, здійснювалися з використанням різних комбінацій параметрів: - DNS-сервер (в табл.1-2, це «*Server name*»), географічне розташування точки вимірювання (в табл.1-2, це «*Location*») та доменне ім'я («*Test domain*»). Це дозволило врахувати просторову варіативність в розміщенні DNS-інфраструктури та мінімізувати вплив випадкових чинників на якість вибірки.

Таблиця 2. Приклад реєстру прецедентів, синтезованого за допомогою ШІ (фрагмент)
Table 2. Example of a registry of precedents synthesized using AI (fragment)

Location	**Server name**	**Test domain**	**Plain query time (ms)**	**DoH query time (ms)**	**DoT query time (ms)**	***Comment***
ISRAEL	Google-Reserve	**nic.ar**	1030	1018	-	*High PQ and DHQ latency*
USA	OpenDNS-Reserve	**post.japanpost.jp**	833	989	1168	*High latency for all protocols*
FRANCE	Google-Reserve	**nic.ar**	231	911	-	*High DHQ latency with low PQ*
ISRAEL	OpenDNS	**gov.za**	76	345	808	*High DTQ latency compared to PQ and DHQ*
FINLAND	OpenDNS	**xinhuanet.com**	430	1156	1316	*High DHQ and DTQ latency compared to PQ.*
FRANCE	Quad9-Reserve	**bbc.co.uk**	296	0	1335	*Zero DHQ latency, also high DTQ.*
ISRAEL	ControlD	**bbc.co.uk**	852	-	861	*PR is null*

Додатковим підтвердженням зазначеної вище, несподіваної логіки висновків ШІ, є коментарі на кшталт *«DTQ = 0»* у випадках, коли інші показники часу затримки DNS запитів виходили за межі норми (приклад див. рис. 2).

Також, варто зазначити, що значна частина коментарів з блоку ШІ, була сфокусована виключно на одному показнику, попри наявність кількох аномальних ознак у межах одного запису. Це свідчить про надмірне зосередження ШІ на окремих екстремальних (*з його точки зору*) параметрах контрольованого процесу.

Anomalous entry:

Location	Server Name	Test Domain	Plain	DoH	DoT
JAPAN	Quad9	terra.com.br	540	261	0

Рис.2 Різниця коментарів (трактувань) системи на алогічні записи
Fig.2 Difference in system comments (interpretations) on anomalous entries

Для усунення наслідків цієї проблеми було запроваджено новий набір інструкцій, які «заохочують» модель приділяти більше уваги загальному контексту подій й уповільнювати формування власних висновків під час обробки кластерів подібних аномалій. Метою цього вдосконалення є: - спонукати ШІ, враховувати більш ширший контекст показників затримки в межах кожного спостереження поточних подій. Тим самим усуваються передумови концентрації ШІ на одному екстремальному значенні, чим зменшується ризик виникнення явища гіперфокусування на окремому аспекті запису та як наслідок, ігнорування інших релевантних даних або формування хибних позитивних результатів/рекомендацій.

Після корегування відповідних інструкцій для блоку ШІ, спостережено помітні зміни у структурі та змісті коментарів до аномальних записів. Так, замість лаконічних односторонніх класифікацій на кшталт «*пікове значення затримки незашифрованого запиту*» або «*час відповіді DoH дорівнює 0*», модель почала формувати більш комплексні (взаємопов'язані) та порівняльні описи. Наприклад, записи, які раніше позначалися лише як: - «*пікове значення затримки незашифрованого запиту*», тепер описувалися у більш конкретизованій формі, такій як: - «*пікове значення затримки незашифрованого запиту, затримка DoH-запиту перебуває в межах норми*». Аналогічно, випадки з нульовими або аномально низькими чи високими значеннями почали супроводжуватися явними порівняннями з іншими показниками затримки, наприклад: «*нульова затримка DoT-запиту за умов високої затримки незашифрованого запиту*» тощо.

Загалом така зміна характеру коментарів моделі, свідчить про ефективність запроваджених заходів, стосовно зменшення ефекту гіперфокусування. Оновлена парадигма логіки повноважень ШІ, демонструє більш цілісну та взаємопов'язану інтерпретацію аномальних випадків затримки DNS-трафіку. Прийняти заходи зменшили ймовірність упередженості висновків ШІ, відносно ролі та місця будь-якого одного з параметрів в межах спостережуваного процесу, сприяючи більш коректному – контекстно-орієнтованому й аргументованому виявленню реальних аномалій. Це підтверджується через таргетовані коментарі, наявність яких покращує умови для проведення інверсного аудиту логіки ШІ *(«**IA logic IA**» – Inverse Audit of the logic of AI)*. В наслідок проведених змін істотно змінилась точність ідентифікації аномалій DNS трафіку для тих записів, які були внесені до реєстру за участю саме модулю ШІ.

В якості прикладу вказаних процесів, нижче (рис. 3) наведено показові діаграми з характерною різницею аномалій контрольованого параметру (*в даному разі, часу затримки запитів)* для наявного переліку доменів реєстру прецедентів в умовах, «ДО» та «ПІСЛЯ» (світло-сіра, *Modified*) проведених корегувань логіки прийняття рішень дослідної моделі ШІ.

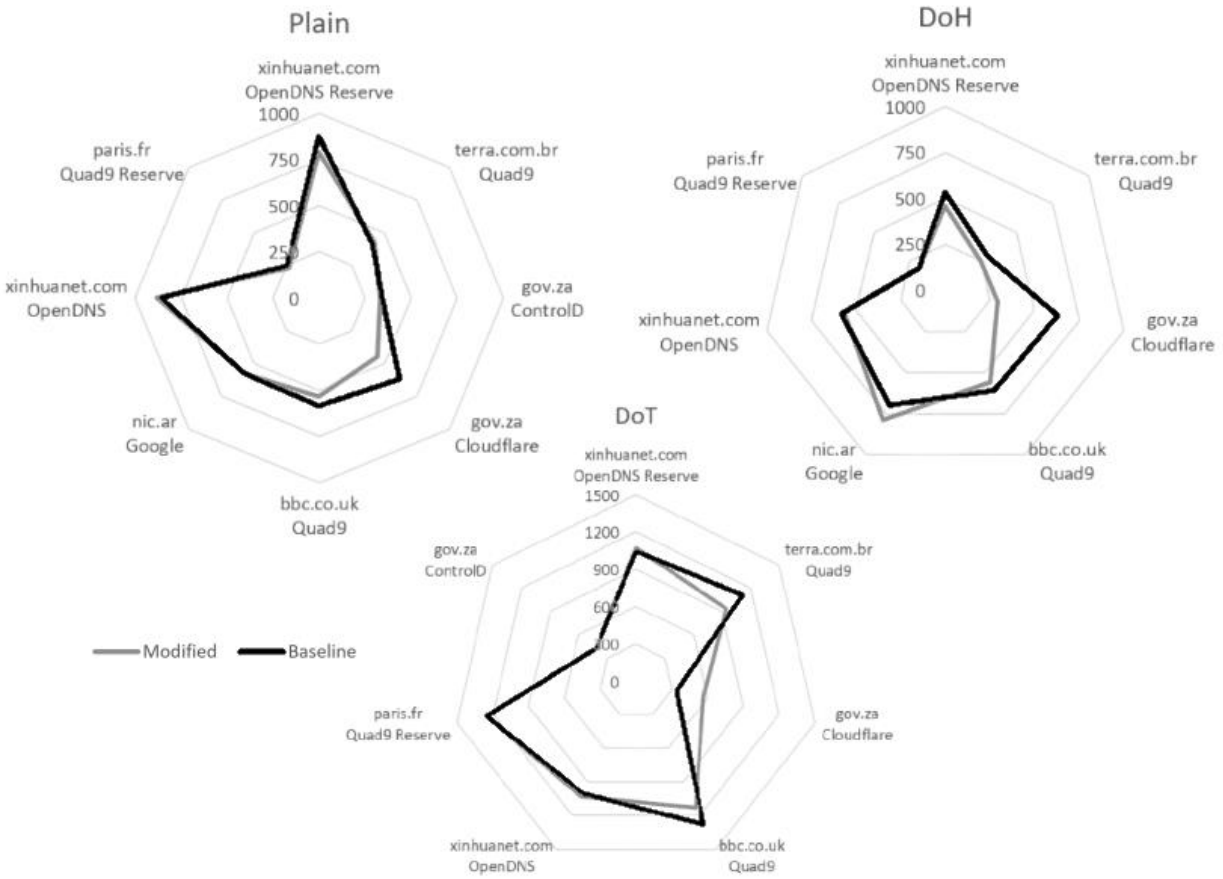


Рис.3 Приклад різниці трактувань аномалій трафіку (шкала часу в [ms])
Fig.3 Example of difference in interpretations of traffic anomalies (time scale in [ms])

Темна/чорна (*Baseline*) лінія на рис. 3 характеризує аномальні показники (середню затримку) «ДО» внесення корегувань. Відповідно, світла лінія характеризує відомості реєстру аномалій, згідно нової парадигмі дій ШІ. Як видно, корегована логіка дій, містить більш ґрунтовний набір даних про виявлені аномалії (за амплітудою часових викидів), що виключає хибні дії/реакції в частині корегування параметрів діючої RPZ. Інше кажучи, результати моніторингу DNS трафіку, стали більш ґрунтовними й виваженими, що відображається через зменшення, як кількості, так й амплітуди «викидів» контрольованого параметра (для умов рис. 3, це часові аномалії DNS запитів). Тобто, чим більш звуженим (чи рівномірним) стає діапазон спостережуваних часових аномалій, тим більш адекватною є реакція системи ШІ на фактичний перебіг контрольованих подій (в т.ч. умов (!) й обставин (!!) процесу, що спостерігається).

4. Висновки

1. Проведено дослідне моделювання програмного інструменту комплексного моніторингу DNS-трафіку із застосуванням концепції прецедентного аналізу (*CBR*), як основи процедурної парадигми ШІ. Підтверджено її придатність для підвищення прозорості виявлення поведінкових аномалій трафіку. Результати моделювання свідчать, що *CBR* підхід, покращує співвіднесення нових спостережень з наявною базою знань, що розширює можливості зворотного аудиту логіки ШІ та підвищує ступінь валідації прийнятих рішень (реакцій ШІ системи).

2. Впровадження *CBR* підвищує прозорість, послідовність і адаптивність штучно синтезованих рішень, та робить поведінкову логіку систем ШІ, більш узгодженою з традиційною логікою мислення людини.

3. Дослідна система успішно доповнила вихідну таблицю прецедентів новими відомостями про аномалії. Підтверджено здатність виявляти, як наперед визначені категорії аномалій, так і додаткові (нові) нерегулярності у даних, що контролюються. Використаний алгоритм обробки

даних забезпечує трактування результатів моніторингу за рамками явно заданих прикладів, водночас зберігаючи/враховуючи логіку причинно-наслідкових зв'язків спостережуваних подій, спираючись на відомості формалізованих шаблонів прецедентів.

4. У ході моделювань визначені обмеження застосованого підходу. Дослідна ШІ система виявила властивість «кластеризації власних трактувань» (ефект т.з. гіперфокусування). Це може призводити до хибної інтерпретації процесів і, як наслідок, до хибних позитивних спрацьовувань. Ефект присутній у випадках, коли подібні відомості розташовані у безпосередній близькості один від одного (кластерах записів). В даному випадку логіка контекстуальній схожості домінує над явними протокольними обмеженнями. В якості компенсаторних заходів, застосовано новий набір обмежень (інструкцій прямої дії). Подальше тестування підтвердило, що ці зміни зменшили прояв виявленого ефекту, підвищивши надійність інтерпретації аномалій.

5. Одержані результати моделювань дозволяють стверджувати, що «логіка роботи систем ШІ не є чимось унікальною і безумовно аксіоматичною». У цьому контексті слід мати на увазі внутрішнє прагнення систем ШІ до самооптимізації в процесі вирішення покладених на них завдань. Це специфічне «прагнення» потребує запровадження додаткових інструкцій та прямих заборон. Запорукою належного виконання подібних обмежувально-керуючих дій є можливість реалізації інверсного аудиту логіки прийнятих ШІ рішень (*«IA logic IA»*).

6. В якості подальших досліджень слід розглядати: - удосконалення механізмів модерації реєстру прецедентів; - масштабування парадигми *CBR* на всі елементи системи хмарного моніторингу DNS трафіку; - розширення можливостей системи, щодо варіативності сценаріїв моніторингу та структури тестових запитів для покращення виявлення нових аномалій.

СПИСОК ЛІТЕРАТУРИ

1. Chepel D., Malakhov S. Multibased cloud monitoring of DNS traffic for operative correction of current RPZ parameters. *Modern information security*. 2025. Т. 63, № 3. С. 176–187. URL: https://doi.org/10.31673/2409-7292.2025.031949 (дата звернення: 23.02.2026).
2. Chepel D., Malakhov S. Summary of DNS traffic filtering trends as a component of modern information systems security. *Computer science and cybersecurity*. 2024. № 1. С. 6–21. URL: https://doi.org/10.26565/2519-2310-2024-1-01 (дата звернення: 23.02.2026).
3. Чепель Д., Малахов С. Мультипротокольний моніторинг трафіку DNS, як основа для коригування поточних параметрів RPZ. *Theoretical and practical aspects of modern scientific research*. 2025. URL: https://doi.org/10.36074/logos-24.01.2025.049 (дата звернення: 23.02.2026).
4. Коробейнікова Т., Федчук Т. Огляд протоколів DNS, DoH та DoT. *Débats scientifiques et orientations prospectives du développement scientifique*. 2024. URL: https://doi.org/10.36074/logos-01.03.2024.056 (дата звернення: 23.02.2026).
5. Advancements in artificial intelligence and data science: models, applications, and challenges / M. F. Safitra et al. *Procedia computer science*. 2024. Т. 234. С. 381–388. URL: https://doi.org/10.1016/j.procs.2024.03.018 (дата звернення: 23.02.2026).
6. Data analysis in the era of generative AI / J. P. Inala et al. 2024. (Препринт). URL: https://doi.org/10.48550/arXiv.2409.18475 (дата звернення: 23.02.2026).
7. Artificial intelligence approaches and mechanisms for big data analytics: a systematic study / A. M. Rahmani et al. *PeerJ computer science*. 2021. Т. 7. e488. URL: https://doi.org/10.7717/peerj-cs.488 (дата звернення: 23.02.2026).
8. Zebin T., Rezvy S., Luo Y. An explainable AI-based intrusion detection system for DNS over HTTPS (DoH) attacks. *IEEE transactions on information forensics and security*. 2022. С. 1. URL: https://doi.org/10.1109/tifs.2022.3183390 (дата звернення: 23.02.2026).
9. Ali B., Chen G. Next-generation AI for advanced threat detection and security enhancement in DNS over HTTPS. *Journal of network and computer applications*. 2025. Т. 244. 104326. URL: https://doi.org/10.1016/j.jnca.2025.104326 (дата звернення: 23.02.2026).
10. Pradeep P., Caro-Martínez M., Wijekoon A. Empowering explainable artificial intelligence through case-based reasoning: a comprehensive exploration. *IEEE transactions on knowledge and data engineering*. 2025. С. 1–20. URL: https://doi.org/10.1109/tkde.2025.3609825 (дата звернення: 23.02.2026).

11. Pradeep P., Caro-Martínez M., Wijekoon A. A practical exploration of the convergence of Case-Based Reasoning and Explainable Artificial Intelligence. *Expert systems with applications*. 2024. 124733. URL: https://doi.org/10.1016/j.eswa.2024.124733 (дата звернення: 23.02.2026).
12. Hatalis K., Christou D., Kondapalli V. Review of case-based reasoning for LLM agents: theoretical foundations, architectural components, and cognitive integration. 2025. (Препринт). URL: https://doi.org/10.48550/arXiv.2504.06943 (дата звернення: 23.02.2026).
13. Гончаров М., Чепель Д., Малахов С. Оцінка обчислювальної складності етапу попередньої обробки вхідних даних при реалізації процедур стегановставки зображень. *Наука і техніка сьогодні*. 2025. № 8(49). URL: https://doi.org/10.52058/2786-6025-2025-8(49)-1228-1245 (дата звернення: 23.02.2026).
14. Горелько М., Малахов С. Аналіз метаданих шифрованого трафіку як чинник нівелювання «сліпих зон» безпеки в сучасних IT - системах. *Інтелектуальні технології у міждисциплінарних дослідженнях: Збірник наукових праць XI МНТК*. Харків: ХНУ ім. В.Н. Каразіна, Україна, 2025. С. 89–92.

REFERENCES

1. D. Chepel and S. Malakhov, “Multibased cloud monitoring of DNS traffic for operative correction of current RPZ parameters,” *Modern Information Security*, vol. 63, no. 3, pp. 176–187, 2025. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.31673/2409-7292.2025.031949
2. D. Chepel and S. Malakhov, “Summary of DNS traffic filtering trends as a component of modern information systems security,” *Computer Science and Cybersecurity*, no. 1, pp. 6–21, Sep. 2024. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.26565/2519-2310-2024-1-01 [in Ukrainian]
3. D. Chepel and S. Malakhov, “Мультипротокольний моніторинг трафіку DNS, як основа для коригування поточних параметрів RPZ [Multi-protocol DNS traffic monitoring as a basis for adjusting current RPZ parameters],” in *Theoretical and Practical Aspects of Modern Scientific Research*. Eur. Scientific Platform, 2025. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.36074/logos-24.01.2025.049 [in Ukrainian]
4. T. Korobeinikova and T. Fedchuk, “Огляд протоколів DNS, DoH та DoT [Overview of DNS, DoH, and DoT protocols],” in *Débats scientifiques et orientations prospectives du développement scientifique*. Eur. Scientific Platform, 2024. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.36074/logos-01.03.2024.056 [in Ukrainian]
5. M. F. Safitra, M. Lubis, T. F. Kusumasari, and D. P. Putri, “Advancements in artificial intelligence and data science: Models, applications, and challenges,” *Procedia Computer Science*, vol. 234, pp. 381–388, 2024. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.1016/j.procs.2024.03.018
6. J. P. Inala *et al.*, *Data Analysis in the Era of Generative AI*. To be published. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.48550/arXiv.2409.18475
7. A. M. Rahmani *et al.*, “Artificial intelligence approaches and mechanisms for big data analytics: A systematic study,” *PeerJ Computer Science*, vol. 7, Apr. 2021, Art. no. e488. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.7717/peerj-cs.488
8. T. Zebin, S. Rezvy, and Y. Luo, “An explainable AI-based intrusion detection system for DNS over HTTPS (DoH) attacks,” *IEEE Transactions on Information Forensics and Security*, p. 1, 2022. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.1109/tifs.2022.3183390
9. B. Ali and G. Chen, “Next-generation AI for advanced threat detection and security enhancement in DNS over HTTPS,” *Journal of Network and Computer Applications*, vol. 244, p. 104326, Dec. 2025. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.1016/j.jnca.2025.104326
10. P. Pradeep, M. Caro-Martínez, and A. Wijekoon, “Empowering explainable artificial intelligence through case-based reasoning: A comprehensive exploration,” *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–20, 2025. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.1109/tkde.2025.3609825
11. P. Pradeep, M. Caro-Martínez, and A. Wijekoon, “A practical exploration of the convergence of Case-Based Reasoning and Explainable Artificial Intelligence,” *Expert Systems With Applications*, p. 124733, Jul. 2024. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.1016/j.eswa.2024.124733

12. K. Hatalis, D. Christou, and V. Kondapalli, *Review of Case-Based Reasoning for LLM Agents: Theoretical Foundations, Architectural Components, and Cognitive Integration*. To be published. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.48550/arXiv.2504.06943
13. М. Гончаров, Д. Чепель, and С. Малахов, "Оцінка обчислювальної складності етапу попередньої обробки вхідних даних при реалізації процедур стегановставки зображень," *Наука і техніка сьогодні*, no. 8(49), Sep. 2025. Accessed: Feb. 23, 2026. [Online]. Available: https://doi.org/10.52058/2786-6025-2025-8(49)-1228-1245 [in Ukrainian]
14. M. Horelko and S. Malakhov, "Аналіз метаданих шифрованого трафіку як чинник нівелювання «сліпих зон» безпеки в сучасних IT – системах [Metadata analysis of encrypted traffic as a factor in eliminating security "blind spots" in modern IT systems]," in *Інтелектуальні технології у міждисциплінарних дослідженнях: Збірник наукових праць XI МНТК*. Харків, Україна: ХНУ ім. В.Н. Каразіна, 2025, pp. 89–92. [in Ukrainian]

Chepel Danylo *Ph.D student*
of the Department of Cybersecurity of Information Systems, Networks and Technologies
V.N. Karazin Kharkiv National University, 4 Svobody sq., Kharkiv, Ukraine, 61022
Malakhov Serhii *Ph.D, Associate Professor*
of the Department of Cybersecurity of Information Systems, Networks and Technologies
V.N. Karazin Kharkiv National University, 4 Svobody sq., Kharkiv, Ukraine, 61022
Honcharov Mykyta *Ph.D student*
of the Department of Cybersecurity of Information Systems, Networks and Technologies
V.N. Karazin Kharkiv National University, 4 Svobody sq., Kharkiv, Ukraine, 61022

Application of a precedent analysis paradigm for the purposes of multibase cloud monitoring of DNS traffic

Relevance. The increasing complexity of DNS infrastructure and the growing level of threats in the network environment necessitate the development of intelligent DNS traffic monitoring tools capable of providing transparent, adaptive, and well-grounded detection of behavioral anomalies. Particular relevance is associated with the implementation of approaches that enhance the traceability of decision-making logic in artificial intelligence (AI) systems.
Purpose. The purpose of this study is to experimentally investigate a prototype software tool for monitoring the current state of DNS traffic with extensive implementation of AI capabilities, the logic of which is based on the concept of case-based reasoning (CBR) for behavioral DNS traffic anomaly analysis.
Research Methods. The study employs simulation modeling methods, multi-base measurements of DNS query processing time using a system of distributed cloud-based sensor-testers, as well as case-based reasoning algorithms for intelligent post-processing of data. The prototype was implemented as a Python client integrated with the Gemini API, operating on a dataset formed based on the results of previous studies [1–2]. During operation, the system autonomously modifies the anomaly registry by adding new cases based on analytical processing results.
Results. The obtained results demonstrate that the proposed DNS traffic monitoring approach ensures the detection of both previously known anomalies and the localization of previously unidentified irregularities. The feasibility of applying the case-based approach to improve the efficiency of adjusting the parameters of the active Response Policy Zone (RPZ) [3] and to enhance situational awareness of personnel regarding DNS traffic security has been confirmed. At the same time, the experiments revealed a so-called "clustering" effect that may lead to false positive event assessments and, consequently, contradictory interpretations of the observed network events.
Conclusions. The revision of existing constraints and analytical tasks for AI modules, followed by further modeling, confirmed that the introduced modifications significantly reduced the identified "clustering" effect and improved the reliability of anomaly interpretation based on a defined system of indirect (implicit) indicators. The obtained results confirm the feasibility of further developing the case-based reasoning approach in intelligent DNS traffic monitoring systems.

Keywords: *information security, artificial intelligence, traffic filtering, DNS, RPZ, CBR, network anomalies, cloud computing, distributed network, DNS protocols.*

Надійшла у першій редакції 25.01.2026; рекомендована до друку 01.03.2026; опублікована 03.04.2026.
The first version has been received on 25.01.2026; was recommended for printing on 01.03.2026; published 03.04.2026.