

МОДЕЛЮВАННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІ Й УПРАВЛІННІ

<https://doi.org/10.26565/2311-2379-2026-110-01>

УДК 004.8:330.43:338.27

А. В. СТАВИЦЬКИЙ *

доктор економічних наук, професор,
професор кафедри економічної кібернетики
ORCID ID: <https://orcid.org/0000-0002-5645-6758>, e-mail: a.stavytskyy@gmail.com

О. М. БАБКОВА *

студентка кафедри економічної кібернетики
ORCID ID: <https://orcid.org/0009-0002-3254-0565>, e-mail: olha.babkova@gmail.com

* Київський національний університет імені Тараса Шевченка,
вул. Володимирська, 60, Київ, 01033, Україна

МАКРОЕКОНОМІЧНЕ ПРОГНОЗУВАННЯ ГЕОПОЛІТИЧНИХ ЗМІН У СВІТІ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

У статті досліджено можливості застосування методів машинного навчання для макроекономічного прогнозування геополітичних змін на основі панельних макроекономічних даних. Актуальність дослідження зумовлена зростанням рівня геополітичної нестабільності у світовій економіці, посиленням міжнародних конфліктів, санкційних обмежень та структурних криз, що суттєво впливають на економічний розвиток країн. У таких умовах підвищується потреба у використанні сучасних аналітичних підходів, здатних забезпечити більш точне прогнозування геополітичних ризиків та виявлення складних нелінійних взаємозв'язків між макроекономічними показниками. Метою дослідження є розробка та реалізація підходу до прогнозування індексу геополітичного ризику із використанням методів машинного навчання. У межах роботи застосовано ансамблеві алгоритми градієнтного бустингу XGBoost та LightGBM, які дозволяють ефективно працювати з високовимірними даними та враховувати складну структуру взаємозв'язків між змінними. Для оцінювання якості прогнозування використано метрики MAE, RMSE та MAPE. Інформаційну базу дослідження сформовано у вигляді панельної структури «країна-рік», що охоплює 43 країни за період 2000-2024 рр. та включає 499 макроекономічних показників, які характеризують виробництво, споживання, зовнішню торгівлю, інституційний розвиток, демографічні та екологічні процеси. Цільовою змінною виступає індекс геополітичного ризику. Для забезпечення коректності прогнозування вибірку поділено на навчальний (2000-2018 рр.) та тестовий (2019-2024 рр.) періоди. Отримані результати свідчать про високу точність прогнозування для обох моделей, значення MAPE не перевищує 1,9%, що підтверджує ефективність застосування ансамблевих алгоритмів машинного навчання у задачах прогнозування геополітичних процесів. Встановлено, що використані моделі демонструють стійкість результатів навіть за наявності пропущених значень та структурних зламів у часових рядах. Додатково застосовано кластерний аналіз, який дозволив виділити групи країн із різним рівнем геополітичної напруженості. У результаті дослідження підтверджено доцільність використання методів машинного навчання для аналізу та прогнозування геополітичних змін. Запропонований підхід може бути використаний як інструмент підтримки прийняття рішень у сфері макроекономічного прогнозування та економічної політики.

Ключові слова: **машинне навчання, геополітичний ризик, макроекономічне прогнозування, ансамблеві алгоритми, класифікація країн.**

JEL Classification: C53, C55, C63, F51, O57.

Постановка проблеми. На сьогодні геополітичні зміни, зокрема міжнародні конфлікти, військові дії, економічні санкції, зміни торговельних режимів, енергетичні кризи та посилення політичної нестабільності, відіграють важливу роль у формуванні макроекономічної динаміки країн та регіонів. В умовах зростання глобальної невизначеності виникає потреба у пошуку нових підходів до аналізу та прогнозування геополітичних процесів. Стандартні методи



економічного аналізу часто виявляються малоефективними у виявленні складних нелінійних залежностей між макроекономічними показниками та геополітичними змінами, що зумовлює перехід до застосування сучасних методів машинного навчання.

Актуальність дослідження пояснюється зростаючою потребою у побудові ефективних прогнозних моделей (International Monetary Fund, 2025), що дозволяють вчасно ідентифікувати потенційні геополітичні зміни на основі макроекономічних даних. Використання методів машинного навчання дозволяє обробляти великі обсяги даних, виявляти складні нелінійні зв'язки між змінними та підвищувати точність прогнозів, що має важливе значення для прийняття управлінських рішень у сфері економічної політики та інвестиційної діяльності.

Аналіз останніх досліджень. Сучасні наукові дослідження свідчать про зростаючу роль геополітичних факторів у формуванні макроекономічної динаміки. Геополітичні зміни розглядаються як трансформації умов міждержавної взаємодії, що виникають унаслідок змін у міжнародних відносинах, зовнішньоекономічній політиці, ролі держави на світовій арені та міжнародних інституцій. У межах світової економіки, яка є системою взаємопов'язаних національних господарств, такі зміни суттєво впливають на рух товарів, капіталу та робочої сили, формуючи зовнішнє середовище економічного розвитку країн (Боярчук та ін., 2018).

У науковій літературі геополітичні процеси часто аналізуються через їх вплив на ключові макроекономічні показники. Зокрема, дослідження демонструють, що геополітичні ризики проявляються у вигляді міжнародних конфліктів, санкцій, змін торговельних режимів і посилення протекціонізму, що безпосередньо впливає на обсяги міжнародної торгівлі, інвестиційні потоки та фінансову стабільність (International Monetary Fund, 2025; Боярчук та ін., 2018). Емпіричні результати, отримані на основі гравітаційних моделей, підтверджують наявність статистично значущого негативного впливу геополітичного ризику на обсяги міжнародної торгівлі, що свідчить про важливість урахування геополітичних факторів у макроекономічному аналізі (Gupta et al., 2019).

Окремий напрям досліджень присвячений аналізу взаємозв'язку між макроекономічною нестабільністю та геополітичними процесами. У звітах міжнародних організацій, зокрема Міжнародного валютного фонду та Світового банку, підкреслюється, що такі показники, як темпи економічного зростання, інфляція, безробіття, валютний курс і стан державних фінансів, можуть виступати як індикаторами, так і факторами геополітичної нестабільності (International Monetary Fund, 2025; World Bank Group, 2025). Водночас у дослідженнях Організації Об'єднаних Націй наголошується на ролі соціально-економічних дисбалансів як потенційних катализаторів конфліктів і політичної дестабілізації (United Nations, 2025).

Важливим аспектом сучасних досліджень є використання індексів геополітичного ризику для кількісної оцінки впливу політичних і міжнародних подій на економічну динаміку. Індекс геополітичного ризику, сформований на основі аналізу текстів міжнародних новинних джерел, дозволяє формалізувати рівень геополітичної напруженості та використовувати його як змінну у макроекономічних моделях (Geopolitical Risk Index, 2018). Емпіричні дослідження показують, що зростання значення цього індексу супроводжується зниженням темпів економічного зростання, підвищенням інфляційних ризиків і скороченням інвестиційної активності (International Monetary Fund, 2025; Caldara & Iacoviello, 2022).

Разом із тим результати досліджень свідчать про складний і нелінійний характер взаємозв'язку між геополітичними ризиками та макроекономічними показниками. На основі агрегованих даних часто не виявляється чітких лінійних залежностей між індексом геополітичного ризику та такими показниками, як інфляція чи інвестиції, що вказує на обмеженість традиційних економетричних підходів у дослідженні геополітичних процесів. Це зумовлює необхідність використання більш гнучких методів аналізу, здатних враховувати багатовимірність і нелінійність відповідних взаємозв'язків.

У зв'язку з цим у сучасній науковій літературі зростає інтерес до застосування методів машинного навчання у макроекономічному прогнозуванні. Дослідження показують, що алгоритми машинного навчання, зокрема градієнтний бустинг, нейронні мережі та ансамблеві методи, забезпечують вищу точність прогнозів порівняно з традиційними економетричними моделями, особливо в умовах нестабільності та структурних змін (Goulet Coulombe et al., 2020). Їх перевагою є здатність працювати з великими обсягами даних, виявляти складні нелінійні залежності та враховувати взаємодію численних факторів.

Окремий напрям досліджень пов'язаний із застосуванням методів машинного навчання для прогнозування політичної та геополітичної нестабільності. Зокрема, моделі машинного навчання використовуються для аналізу макроекономічних, соціальних і політичних індикаторів з метою виявлення зон підвищеного ризику конфліктів і кризових явищ. Отримані результати свідчать про їх здатність забезпечувати більш точні прогнози та виступати інструментом раннього попередження геополітичних ризиків (Хуе et al., 2025).

Важливим доповненням до кількісних макроекономічних даних є використання текстових джерел інформації. Методи обробки природної мови дозволяють аналізувати новинні потоки та аналітичні матеріали, формуючи індикатори геополітичного ризику, які можуть бути інтегровані у моделі машинного навчання. Поєднання структурованих макроекономічних даних і неструктурованої текстової інформації розглядається як перспективний напрям підвищення точності прогнозування макроекономічних наслідків геополітичних змін (Caldara & Iacoviello, 2022).

Таким чином, аналіз наукової літератури показав, що геополітичні зміни є важливим фактором макроекономічної динаміки, а їх дослідження потребує використання сучасних аналітичних підходів. Водночас у наявних дослідженнях недостатньо уваги приділено використанню методів машинного навчання для прогнозування індексу геополітичного ризику на основі великомасштабних панельних макроекономічних даних. Більшість робіт зосереджена або на аналізі впливу геополітичних ризиків на окремі макроекономічні показники, або на прогнозуванні традиційних економічних змінних, таких як ВВП чи інфляція. При цьому питання застосування ансамблевих алгоритмів машинного навчання для прогнозування геополітичних змін у багатofакторному міжнародному середовищі залишається недостатньо дослідженим.

У межах даного дослідження висувається дослідницька гіпотеза про те, що ансамблеві алгоритми машинного навчання здатні забезпечувати високу точність прогнозування індексу геополітичного ризику на основі макроекономічних показників та ефективно враховувати складні нелінійні взаємозв'язки між змінними. Додатково передбачається, що використання високимірних панельних даних сприяє підвищенню якості прогнозування геополітичних змін завдяки можливості врахування значної кількості макроекономічних факторів та їх взаємозв'язків.

Мета та завдання. Метою дослідження є розробка та застосування моделей машинного навчання для прогнозування геополітичних змін на основі макроекономічних показників.

Для досягнення поставленої мети передбачено розв'язання таких завдань:

- дослідити вплив геополітичних змін на макроекономічні процеси;
- визначити макроекономічні показники, що можуть використовуватися для прогнозування геополітичного ризику;
- обґрунтувати доцільність застосування методів машинного навчання для прогнозування геополітичних змін;
- сформувати інформаційну базу дослідження та побудувати моделі машинного навчання для прогнозування індексу геополітичного ризику;
- оцінити результати прогнозування та визначити ключові фактори формування геополітичного ризику;
- сформулювати практичні висновки щодо використання методів машинного навчання у макроекономічному прогнозуванні.

Методологія дослідження.

Прогнозування. Макроекономічне прогнозування геополітичних змін реалізовано за допомогою ансамблевих алгоритмів машинного навчання XGBoost та LightGBM, що належать до класу градієнтного бустингу на основі дерев рішень. Використання саме цих алгоритмів зумовлене специфікою поставленої задачі: прогнозування індексу геополітичного ризику (GPR) здійснюється на основі великомасштабної панельної бази даних, яка включає 499 макроекономічних показників для 43 країн за період 2000-2024 рр. (Caldara & Iacoviello, 2022). Така висока розмірність простору ознак, наявність потенційно складних нелінійних взаємозв'язків, мультиколінеарності та структурних зламів обмежує застосування класичних лінійних регресійних моделей, тоді як ансамблеві методи бустингу дозволяють послідовно мінімізувати функцію втрат шляхом побудови композиції слабких моделей (дерев рішень), кожне з яких уточнює помилки попередніх (Mullainathan & Spiess, 2017).

Алгоритми XGBoost та LightGBM належать до класу градієнтного бустингу дерев рішень і реалізують адитивну модель, у межах якої прогноз формується як сума базових функцій:

$$\hat{y}_{i,t} = \sum_{k=1}^K \alpha \cdot f_k(X_{i,t}) \quad (1)$$

де f_k – k -те дерево рішень; K – загальна кількість ітерацій (дерев); α – коефіцієнт навчання, що масштабує внесок кожного дерева (Chen & Guestrin, 2016; Ke et al., 2017).

Навчання здійснюється поетапно. На кожній ітерації k додається нове дерево f_k , що мінімізує функцію втрат відносно поточних залишків. Для регресійної задачі використовується квадратична функція втрат (Chen & Guestrin, 2016):

$$l(y_{i,t}, \hat{y}_{i,t}) = (y_{i,t} - \hat{y}_{i,t})^2 \quad (2)$$

У XGBoost задача оптимізації формулюється у регуляризованому вигляді:

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (3)$$

де $\Omega(f_k)$ – штраф за складність дерева, що залежить від кількості листків та величин wig листка (L1 та L2 регуляризація) (Chen & Guestrin, 2016).

LightGBM реалізує аналогічну адитивну схему мінімізації втрат, проте відрізняється стратегією побудови дерев. Замість симетричного росту по рівнях застосовується leaf-wise підхід, за якого на кожному кроці розширюється листок із максимальним зменшенням функції втрат. Така стратегія дозволяє швидше досягати локального мінімуму при роботі з великими вибірками. Контроль складності здійснюється через обмеження на глибину дерева, кількість листків та мінімальну кількість спостережень у листку (Ke et al., 2017).

Отже, обидва алгоритми реалізують послідовну мінімізацію функції втрат шляхом додавання дерев до ансамблю, але відрізняються механізмом оптимізації та процедурою росту структури дерева, що й обґрунтовує їх паралельне застосування для перевірки стійкості результатів.

Інформаційна база дослідження сформована у вигляді панельної структури «країна-рік». Для кожного спостереження фіксується значення цільової змінної – індексу геополітичного ризику (Geopolitical Risk Index, 2018) – та вектор із 499 макроекономічних показників (World Bank Group, 2024), що охоплюють характеристики виробництва, споживання, заощаджень, зовнішньої торгівлі, екологічного навантаження, демографічних процесів та інституційних параметрів розвитку. Загальний період спостережень становить 2000-2024 рр., що дозволяє охопити як відносно стабільні фази світової економіки, так і періоди кризових зламів.

З метою забезпечення коректності прогнозової оцінки вибірку поділено хронологічно: навчальна вибірка охоплює 2000-2018 рр., а тестова – 2019-2024 рр. Такий підхід виключає витік інформації з майбутніх періодів у процес навчання та імітує реальні умови прогнозування. Вибір саме 2019-2024 рр. як тестового періоду є принципово важливим, оскільки він включає фазу глобальних геополітичних шоків, зокрема різке зростання ризику у 2022 році, що дозволяє оцінити здатність моделей відтворювати структурні злами та екстремальні коливання (Caldara et al., 2025).

Оскільки у вихідній базі даних наявні пропущені значення, дослідження проведено у двох режимах: Raw data (без заповнення пропусків) та Filled data (з попередньо заповненими пропусками). Це дозволяє оцінити чутливість результатів до способу попередньої обробки даних та перевірити робастність висновків.

Реалізація моделей виконана мовою Python із використанням стандартних етапів підготовки даних для підвищення стійкості та точності прогнозування. Для цільової змінної застосовано логарифмічне перетворення, що зменшує вплив екстремальних значень і стабілізує дисперсію помилок. Для врахування часової інерційності геополітичних процесів додано лагову змінну lag_1 (значення GPR у попередньому році) та показник risk_momentum, який відображає попередню зміну індексу.

Окремо введено змінну `wag_dummu` для України: у 2014-2021 рр. вона дорівнювала 1, а у 2022 році – 10 як індикатор різкого посилення геополітичного шоку через повномасштабне вторгнення. Для інших країн і для України до 2014 року значення становило 0. Це дозволяє моделі краще враховувати переломні етапи динаміки ризику в українському випадку.

Додатково використано ваги спостережень (`sample_weight`), зокрема посилене зважування 2022 року для точнішого відтворення пікових значень GPR у період найбільшого геополітичного шоку.

Навчання моделей здійснювалося з використанням фіксованих гіперпараметрів, підібраних з урахуванням високої розмірності ознак, панельної структури даних і необхідності збалансувати гнучкість моделі та ризик перенавчання.

Для моделі XGBoost використано такі параметри: кількість дерев (`n_estimators=1000`), коефіцієнт навчання (`learning_rate=0.04`), максимальна глибина дерева (`max_depth=6`), параметри L1- та L2-регуляризації (`reg_alpha=1`, `reg_lambda=15`), частка спостережень для побудови дерева (`subsample=0.8`) та частка ознак для кожного дерева (`colsample_bytree=0.8`). Така конфігурація забезпечує виявлення складних нелінійних залежностей, зменшує ризик перенавчання та підвищує стійкість до шуму. Для відтворюваності результатів використано фіксоване початкове значення генератора випадкових чисел (`random_state=42`).

Для моделі LightGBM застосовано аналогічну кількість дерев (`n_estimators=1000`) і коефіцієнт навчання (`learning_rate=0.04`). Також використано максимальну глибину дерева (`max_depth=7`), максимальну кількість листків (`num_leaves=31`), мінімальну кількість спостережень у листку (`min_child_samples=20`) та параметри регуляризації (`reg_lambda=15`, `reg_alpha=2`). Це дозволяє контролювати складність моделі, запобігати нестійким розбиттям і стабілізувати оцінки в умовах можливої мультиколінеарності.

Таким чином, архітектура моделей поєднує велику кількість слабких дерев, невеликий крок навчання та посилену регуляризацію, що забезпечує високу точність прогнозування при збереженні узагальнювальної здатності на тестовій вибірці.

Для оцінювання якості прогнозування використано метрики MAE, RMSE та MAPE.

Аналіз впливу факторів здійснювався на основі показника Gain. Він відображає зменшення функції втрат різниці розбиття вузла дерева за певною ознакою та формально може бути представлений як різниця між значенням функції втрат до поділу та сумою значень після поділу:

$$\text{Gain} = L(\text{parent}) - L(\text{left}) - L(\text{right}) \quad (4)$$

де L – значення функції втрат (Chen & Guestrin, 2016). Підсумковий Gain для кожної ознаки визначається як сума приростів по всіх деревах ансамблю. Чим вищий Gain, тим більший внесок змінної у покращення прогнозної точності моделі. Важливо підкреслити, що оцінка важливості за Gain відображає внесок у прогноз, а не причинно-наслідковий ефект.

Класифікація країн. З метою доповнення результатів прогнозування доцільним є здійснення класифікації країн за рівнем геополітичного ризику. На відміну від задач регресійного прогнозування, у цьому випадку відсутня попередньо визначена інформація про належність країн до певних категорій ризику. Оскільки міжнародні статистичні джерела (World Bank Group, 2023; World Bank Group, 2024; OECD, 2024) не містять формалізованого поділу держав за рівнем геополітичного ризику, виникає потреба у використанні методів навчання без учителя (James et al., 2021).

Серед основних підходів до кластеризації виділяють ієрархічні методи, алгоритми на основі щільності (DBSCAN), спектральну кластеризацію та метод k-means (James et al., 2021; Huang et al., 2025). Вибір алгоритму залежить від структури даних, обсягу вибірки та мети дослідження.

Ієрархічна кластеризація не потребує попереднього визначення кількості кластерів, однак її результати суттєво залежать від вибору метрики відстані та способу агрегації, що може ускладнювати відтворюваність і економічну інтерпретацію результатів (James et al., 2021).

Алгоритми на основі щільності, зокрема DBSCAN, ефективні для виявлення кластерів довільної форми та шумових спостережень, проте потребують наявності виражених зон різної щільності та налаштування параметрів (Huang et al., 2025). У межах даного дослідження вибірка з 43 країн і набір агрегованих ознак є відносно однорідними, що знижує доцільність використання таких підходів.

З огляду на це у дослідженні обрано метод k-means, який дозволяє формувати компактні та внутрішньо однорідні групи за мінімізації внутрішньокластерної дисперсії (James et al., 2021). Важливою перевагою методу є висока інтерпретованість: центр кожного кластера відображає середній профіль ознак відповідної групи країн, що дає змогу трактувати кластери як типові моделі геополітичного ризику. Крім того, k-means є одним із найбільш поширених методів багатовимірної кластеризації у соціально-економічних дослідженнях (Gupta et al., 2023; James et al., 2021).

Базовим показником, що використовується для формування простору ознак, є індекс геополітичного ризику (GPR) (Geopolitical Risk Index, 2018). Індекс базується на текстовому аналізі міжнародних ЗМІ та відображає інтенсивність згадувань воєн і конфліктів. Емпіричні дослідження підтверджують його вплив на макроекономічну динаміку, інфляційні процеси та фінансові умови (Caldara et al., 2025; Bouri et al., 2023; Zhang et al., 2024).

Для формування класифікації використано період 2019-2024 рр., що відповідає тестовій вибірці прогнозної моделі. Даний часовий інтервал характеризується підвищеною глобальною нестабільністю (пандемія COVID-19, енергетичні кризи, повномасштабна війна), що дозволяє класифікувати країни за сучасною конфігурацією геополітичних ризиків.

Для кожної країни формується вектор агрегованих характеристик геополітичного ризику:

$$z_i = (G\bar{P}R_i, \max(GPR)_i, \sigma(GPR)_i, \beta_i, \Delta_i, m_i) \quad (5)$$

де $G\bar{P}R_i$ – середній рівень індексу геополітичного ризику за 2019-2024 рр.; $\max(GPR)_i$ – максимальне (пікове) значення індексу за період; $\sigma(GPR)_i$ – стандартне відхилення, що характеризує волатильність ризику; β_i – коефіцієнт тренду, який відображає середню швидкість зміни ризику у часі; Δ_i – величина шоку 2022 року (відхилення від середнього рівня 2019-2021 рр.); m_i – середня річна зміна індексу за досліджуваний період.

Формування вектора ознак ґрунтується на класичних підходах до опису часових рядів у прикладній економетриці та методах конструювання ознак у машинному навчанні (James et al., 2021; Hyndman & Athanassopoulos, 2021), що забезпечує статистичну обґрунтованість та економічну інтерпретованість кожного показника.

Таким чином, вектор z_i одночасно враховує рівень, нестабільність, довгострокову тенденцію та реакцію країни на екстремальні геополітичні шоки.

Оскільки ознаки мають різні масштаби вимірювання, перед застосуванням алгоритму здійснюється стандартизація.

Алгоритм k-means формалізується як задача мінімізації внутрішньокластерної суми квадратів відхилень:

$$\min_{\{\mu_k\}_{k=1}^K} \sum_{i=1}^N \min_k \|z_i^* - \mu_k\|^2 \quad (6)$$

де μ_k – центр k-го кластера; K – кількість кластерів (James et al., 2021).

Кількість кластерів визначається на основі методу «лікоть» (James et al., 2021) та індексу силуєту (Rousseeuw, 1987).

Для забезпечення відтворюваності результатів кластеризації в алгоритмі k-means зафіксовано параметр `random_state=42`, що гарантує однакову ініціалізацію центрів кластерів при кожному запуску. Додатково використано параметр `n_init=50`, який передбачає багаторазову ініціалізацію з вибором рішення з мінімальною внутрішньокластерною дисперсією, що підвищує стабільність і надійність класифікації.

Метод «лікоть» базується на аналізі внутрішньокластерної суми квадратів відхилень:

$$WCSS(K) = \sum_{k=1}^K \sum_{i \in C_k} \|z_i^* - \mu_k\|^2 \quad (7)$$

де C_k – множина об'єктів k-го кластера; μ_k – центр кластера; z_i^* – стандартизований вектор ознак країни (James et al., 2021).

Із зростанням K значення WCSS монотонно зменшується, однак після певного моменту темп зменшення суттєво сповільнюється. Точка перегину графіка інтерпретується як оптимальна кількість кластерів, оскільки подальше дроблення груп не забезпечує істотного зниження внутрішньої дисперсії.

Для додаткової перевірки використано індекс силуету, який оцінює ступінь відокремленості кластерів. Для кожної країни i обчислюється:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (8)$$

де $a(i)$ – середня відстань від країни до інших об'єктів свого кластера; $b(i)$ – мінімальна середня відстань до об'єктів найближчого іншого кластера (Rousseeuw, 1987).

Значення $s(i)$ належить інтервалу $[-1, 1]$. Чим ближче середнє значення індексу силуету до 1, тим чіткіше відокремлені кластери. Низькі або від'ємні значення свідчать про слабе розмежування груп.

Таким чином, застосування методу k-means до агрегованих характеристик індексу GPR дозволяє сформувати аналітично обґрунтовану класифікацію країн за рівнем ризику геополітичних змін, що доповнює результати прогнозування.

Основні результати дослідження.

Прогнозування. Результати оцінювання прогнозовної якості для кожної моделі та варіанту інформаційної бази наведено у табл. 1.

Таблиця 1 – Порівняльна характеристика результатів прогнозування алгоритмами XGBoost та LightGBM

Table 1 – Comparative characteristics of prediction results using XGBoost and LightGBM algorithms

Алгоритм	База даних	MAE	RMSE	MAPE, %
XGBoost	Raw data	0,0068	0,0155	1,8566
	Filled data	0,0067	0,0154	1,8889
LightGBM	Raw data	0,0067	0,0156	1,8629
	Filled data	0,0068	0,0158	1,9055

Джерело: розраховано автором / Source: calculated by the author

Отримані результати свідчать про високу точність прогнозування для обох алгоритмів: значення MAE перебувають у межах 0,0067-0,0068, RMSE – 0,0154-0,0158, а MAPE не перевищує 1,9 %. Відмінності між XGBoost і LightGBM є незначними, однак найкраще значення MAPE демонструє XGBoost (1,8566 %). Заповнення пропусків не забезпечує суттєвого покращення якості прогнозу, що свідчить про здатність ансамблевих алгоритмів ефективно працювати навіть із неповними даними.

Робастність результатів перевірялася шляхом використання двох незалежних алгоритмів, двох варіантів підготовки даних (Raw та Filled data) та відкладеної тестової вибірки 2019-2024 рр. Стабільність метрик у всіх комбінаціях свідчить про відсутність суттєвого перенавчання та підтверджує узагальнювальну здатність моделей.

Окрім базового хронологічного поділу (2000-2018 / 2019-2024 рр.), додатково було здійснено перевірку моделей на альтернативних часових розбивках вибірки (зміщення межі навчального та тестового періодів). Зокрема, тестування проводилося з використанням скорочених навчальних періодів та відповідно змінених тестових відрізків. У всіх варіантах значення MAPE залишалися меншим за 2 %, а прогнозні траєкторії на графіках зберігали високу узгодженість із фактичними значеннями індексу.

На рис. 1 зображено динаміку середнього фактичного та прогнозного значення індексу геополітичного ризику у тестовому періоді 2019-2024 рр. для моделі XGBoost, побудованої на базі даних без заповнення пропусків. Візуально спостерігається майже повне накладання прогнозної та фактичної кривих у 2019-2024 рр., що підтверджує високу узгодженість моделі з

реальними значеннями та її здатність відтворювати як загальний тренд геополітичного ризику, так і переломні моменти динаміки. У 2019-2021 рр. індекс характеризується помірними значеннями та незначними коливаннями, тоді як у 2022 р. модель адекватно фіксує різкий стрибок GPR, пов'язаний із масштабним геополітичним шоком. У 2023 р. спостерігається часткове зниження індексу, а у 2024 р. – помірне зростання. Загалом XGBoost у режимі Raw data забезпечує адекватне прогнозування середніх річних значень індексу GPR по вибірці.

Аналогічна динаміка спостерігається і для інших конфігурацій моделей (XGBoost – Filled data, LightGBM – Raw data, LightGBM – Filled data), що підтверджує стійкість отриманих результатів.

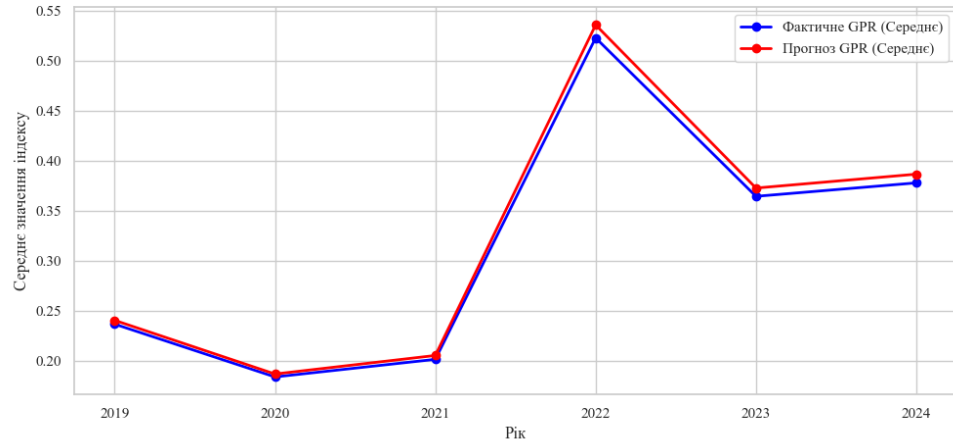


Рис. 1. Динаміка фактичних та прогнозних значень індексу геополітичного ризику (модель XGBoost, Raw data)

Fig. 1. Dynamics of actual and forecast values of the geopolitical risk index (XGBoost model, Raw data)

Джерело: побудовано автором / Source: constructed by the author

На рис. 2 представлено топ-10 макроекономічних факторів, які мають найбільший внесок у формування прогнозу індексу геополітичного ризику за результатами моделі XGBoost на основі бази даних без заповнення пропусків.

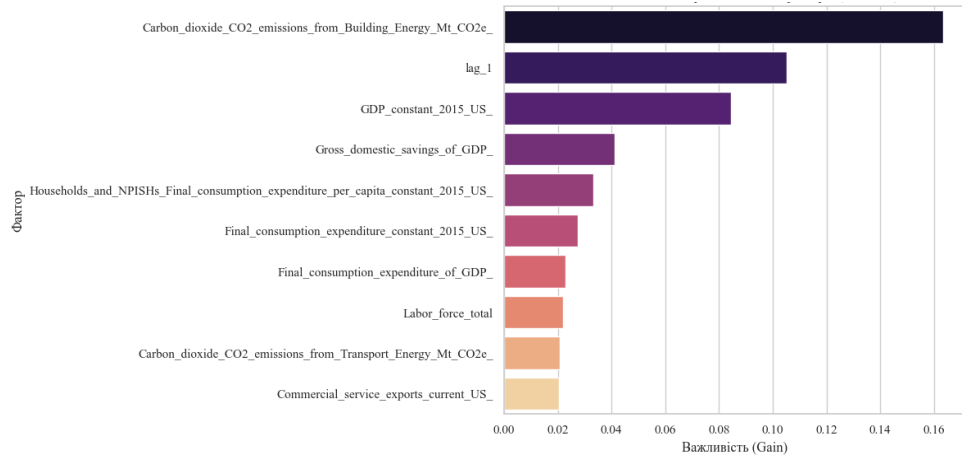


Рис. 2. Рейтинг найважливіших макроекономічних факторів (XGBoost, Raw data)

Джерело: побудовано автором / Source: constructed by the author

Домінування викидів CO₂ від енергоспоживання будівель відображає роль енергетичної інтенсивності економіки та її вразливість до енергетичних шоків. Висока важливість lag_1 підтверджує інерційність ризику, тобто суттєвий вплив попереднього рівня напруженості на поточний. Значущими також є реальний ВВП, валові заощадження та споживчі витрати на душу населення, які характеризують масштаби економіки, її фінансову стійкість і внутрішню соціально-економічну стабільність.

На рис. 3 наведено топ-10 макроекономічних факторів за важливістю для моделі XGBoost у випадку використання заповненої бази даних.

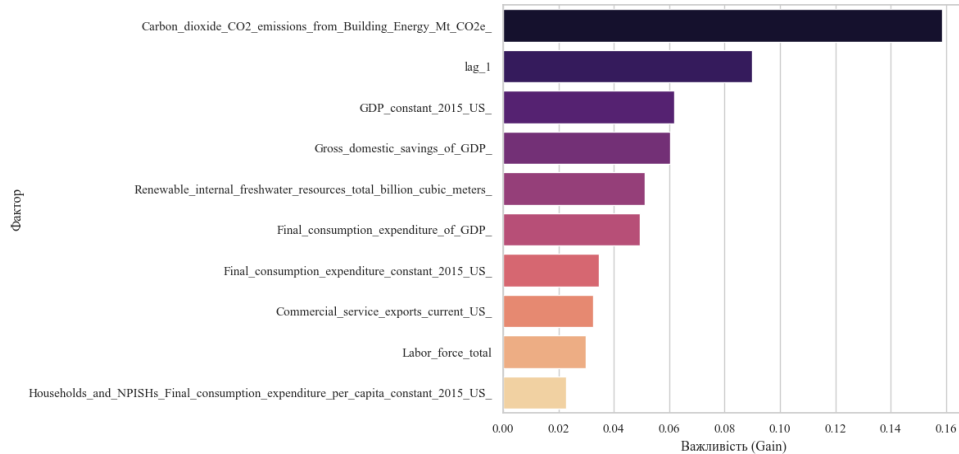


Рис. 3. Рейтинг найважливіших макроекономічних факторів (XGBoost, Filled data)
Fig. 3. Ranking of the most important macroeconomic factors (XGBoost, Filled data)

Джерело: побудовано автором / Source: constructed by the author

Склад топ-5 загалом зберігається, однак замість споживчих витрат до нього входить показник відновлюваних внутрішніх ресурсів прісної води. Це свідчить про посилення ролі екологічно-ресурсних факторів після заповнення пропусків і підкреслює значення ресурсної забезпеченості для геополітичної та економічної вразливості.

Рис. 4 відображає топ-10 макроекономічних факторів за показником важливості, визначених алгоритмом LightGBM на основі незаповненої бази даних.

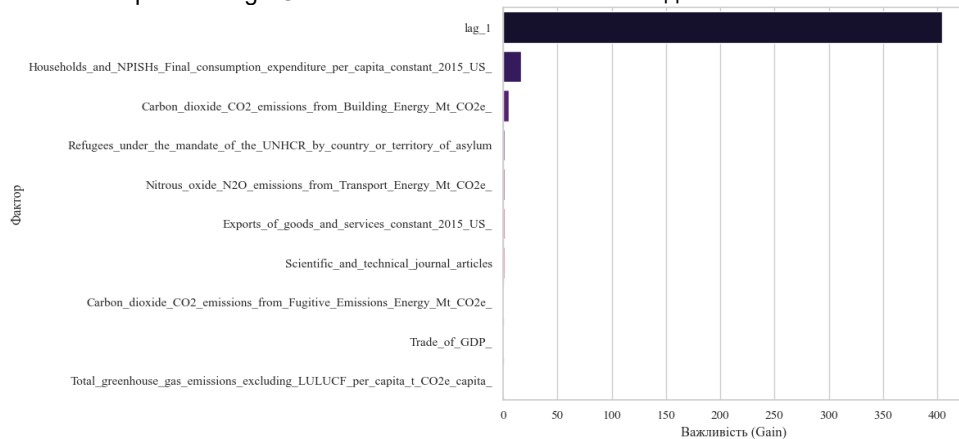


Рис. 4. Рейтинг найважливіших макроекономічних факторів (LightGBM, Raw data)
Fig. 4. Ranking of the most important macroeconomic factors (LightGBM, Raw data)

Джерело: побудовано автором / Source: constructed by the author

Особливістю моделі є виражене домінування змінної lag_1 , що свідчить про сильну інерційність геополітичного ризику та залежність поточного рівня ризику від попереднього. Водночас така концентрація на одній змінній може зменшувати чутливість моделі до структурних змін. Серед інших важливих факторів виділяються показники споживання домогосподарств, викидів CO_2 , зовнішньої торгівлі, науково-технічної активності та міграційних процесів, які відображають соціально-економічні канали поширення ризиків.

На рис. 5 наведено рейтинг найважливіших макроекономічних факторів для моделі LightGBM на основі заповненої бази даних.

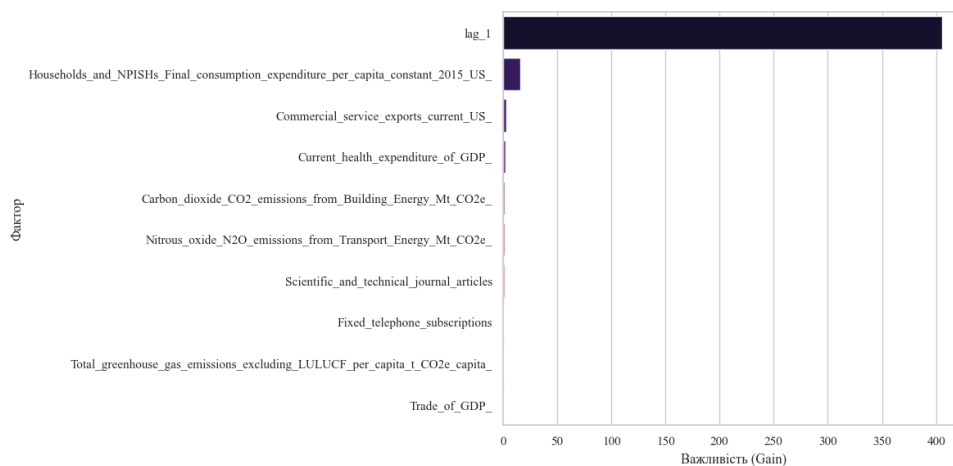


Рис. 5. Рейтинг найважливіших макроекономічних факторів (LightGBM, Filled data)
Fig. 5. Ranking of the most important macroeconomic factors (LightGBM, Filled data)

Джерело: побудовано автором / Source: constructed by the author

Як і у випадку Raw data, домінуючою ознакою залишається lag_1 , що підтверджує ключову роль інерційності геополітичного ризику. Серед інших важливих факторів виділяються споживчі витрати домогосподарств, експорт комерційних послуг, витрати на охорону здоров'я, викиди CO_2 та показники технологічної інфраструктури.

Економічно це означає, що геополітична напруга має накопичувальний характер, тоді як добробут, відкритість економіки, інституційна спроможність і ресурсно-енергетична структура виступають ключовими каналами формування та посилення ризиків.

Результати емпіричного дослідження підтвердили високу ефективність ансамблевих алгоритмів XGBoost та LightGBM для прогнозування індексу геополітичного ризику на основі панельних макроекономічних даних. Обидві моделі забезпечили високу точність на тестовій вибірці 2019-2024 рр., включно з періодом різкого зростання ризику у 2022 році, що свідчить про їхню здатність відтворювати як тренд, так і структурні зміни.

Найкращі результати за сукупністю метрик продемонструвала модель XGBoost на Raw data (мінімальний MAPE та нижчий RMSE). Заповнення пропусків не дало суттєвого покращення, тому використання Raw data є обґрунтованим.

Аналіз важливості факторів показав, що геополітичний ризик має виражену інерційність (lag_1), а також формується під впливом енергетичних, макрофінансових, ресурсних і соціально-економічних чинників (викиди CO_2 , ВВП, заощадження, споживчі витрати тощо). Це підтверджує багатофакторну природу геополітичної нестабільності.

Таким чином, базовою моделлю для практичного застосування доцільно вважати XGBoost, яка забезпечує оптимальне поєднання точності, стійкості та економічної інтерпретованості результатів.

З економічної точки зору результати є змістовними: інерційність ризику відображає його накопичувальний характер і повільне згасання у часі. Значущість макроекономічних показників (ВВП, заощадження, споживчі витрати, показники зовнішньої торгівлі, енергетичні та екологічні індикатори) узгоджується з впливом структури економіки, рівня відкритості, ресурсної

залежності та фінансової стійкості. Енергетичні та ресурсні фактори можуть виступати каналами поширення міжнародних шоків і конфліктів, тоді як макрофінансові показники відображають адаптаційні можливості економіки. Отже, виявлені закономірності відповідають логіці функціонування сучасної глобальної економіки та мають практичну цінність для аналізу і прогнозування геополітичних ризиків.

Класифікація країн. Отже, емпірична реалізація запропонованої методики дозволила здійснити кластеризацію 43 країн за агрегованими характеристиками індексу геополітичного ризику у 2019-2024 рр. Результати визначення оптимальної кількості кластерів на основі методу «лікоть» та індексу силуету наведено у табл. 2 та на рис. 6-7.

Таблиця 2 – Обґрунтування кількості кластерів
Table 2 – Justification of the number of clusters

К	WCSS	Silhouette
2	111,0285	0,2702
3	75,2534	0,2074
4	49,7126	0,1762
5	39,6258	0,1928
6	31,1861	0,1821

Джерело: розраховано автором / Source: calculated by the author

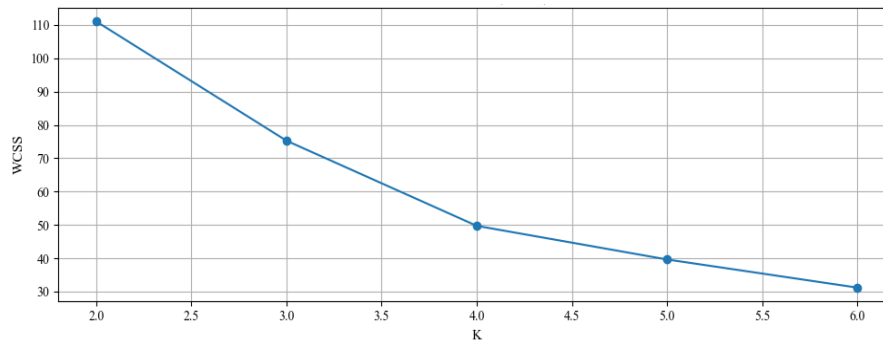


Рис. 6. Обґрунтування кількості кластерів за допомогою методу «лікоть»
Fig. 6. Justification of the number of clusters using the “elbow” method

Джерело: побудовано автором / Source: constructed by the author

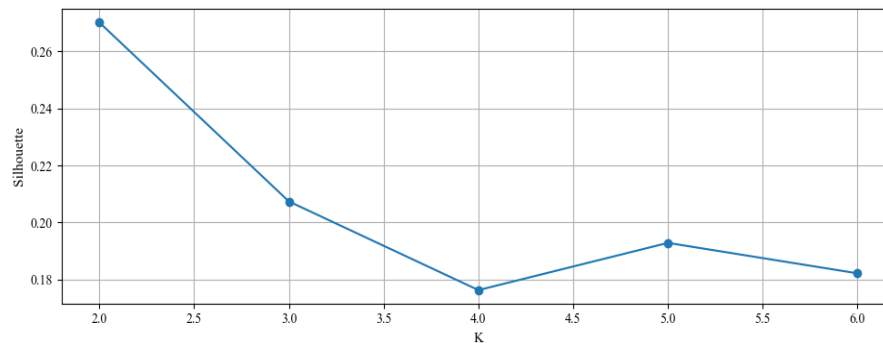


Рис. 7. Обґрунтування кількості кластерів за допомогою індексу силуету
Fig. 7. Justification of the number of clusters using the silhouette index

Джерело: побудовано автором / Source: constructed by the author

Аналіз значень внутрішньокластерної суми квадратів відхилень (WCSS) свідчить про істотне зниження показника при переході від $K=2$ до $K=3$ та подальше зменшення при $K=4$. Після значення $K=4$ темп зниження WCSS помітно сповільнюється, що вказує на наявність характерного «перегину» графіка в області чотирьох кластерів. Це відповідає критерію методу «лікоть» і свідчить про доцільність вибору $K=4$ як оптимальної кількості кластерів.

Значення індексу силуету для досліджуваних варіантів кластеризації знаходяться у діапазоні від 0,176 до 0,270. Найвище значення показника спостерігається при $K=2$ (0,270), однак зі збільшенням кількості кластерів значення індексу поступово зменшується. При $K=4$ індекс силуету становить 0,176, що вказує на прийнятний рівень відокремленості кластерів з урахуванням значної варіації показників геополітичного ризику між країнами.

Таким чином, оптимальною кількістю кластерів у межах даного дослідження обрано $K=4$.

Подальша інтерпретація результатів ґрунтується на аналізі центрів кластерів у реальних одиницях вимірювання (табл. 3).

Отримані центри кластерів демонструють чітку та послідовну градацію країн за основними характеристиками геополітичного ризику. Порівняння середніх значень показників у кожному кластері дозволяє виділити чотири типи країн, що відрізняються за рівнем, волатильністю та динамікою геополітичної напруженості. Кожен кластер відображає певний тип геополітичного середовища, у якому функціонують відповідні країни.

Таблиця 3 – Характеристика центрів кластерів
Table 3 – Characteristics of cluster centers

Група ризику	$G\bar{P}R_i$	$\max(GPR)_i$	$\sigma(GPR)_i$	β_i	Δ_i	m_i
Дуже низький	0,0351	0,0511	0,0115	0,0017	0,0084	-0,0007
Низький	0,1080	0,1928	0,0447	0,0136	0,1138	0,0080
Середній	0,2071	0,3105	0,0607	-0,0051	0,0487	-0,0154
Високий	0,8565	1,6665	0,4296	0,1474	0,9713	0,1021

Джерело: розраховано автором / Source: calculated by the author

Кластер дуже низького ризику характеризується мінімальними значеннями середнього рівня індексу геополітичного ризику ($G\bar{P}R_i$) та його максимального значення ($\max(GPR)_i$), а також дуже низькою волатильністю ($\sigma(GPR)_i$). Реакція на геополітичний шок 2022 року є незначною (Δ_i), а середня річна зміна індексу практично відсутня (m_i). Це означає, що країни цього кластеру функціонують у стабільному та передбачуваному геополітичному середовищі з мінімальними зовнішньополітичними загрозами, де ризики не мають суттєвого впливу на економічну стабільність.

Кластер низького ризику характеризується дещо вищими значеннями середнього ($G\bar{P}R_i$) максимального рівня індексу ($\max(GPR)_i$), а також більшою волатильністю ($\sigma(GPR)_i$). Для цієї групи характерний позитивний тренд зміни індексу (β_i) та помітніший вплив геополітичного шоку 2022 року (Δ_i). Це свідчить про наявність окремих періодів зростання геополітичної напруженості, однак загалом середовище залишається відносно стабільним, а ризики не набувають системного характеру.

Кластер середнього ризику характеризується подальшим зростанням середнього ($G\bar{P}R_i$) та максимального значення індексу ($\max(GPR)_i$), а також вищою волатильністю ($\sigma(GPR)_i$). Для цієї групи характерна більш нестабільна динаміка геополітичного ризику: коефіцієнт тренду є дещо від'ємним (β_i), що може свідчити про поступове зниження ризику після періодів загострення. Водночас середня річна зміна індексу також має від'ємне значення (m_i), що відображає певну стабілізацію ситуації після попередніх коливань. Економічно цей кластер відображає країни, у яких геополітичні ризики виникають регулярніше та створюють додаткову невизначеність для розвитку та міжнародної взаємодії.

Кластер високого ризику демонструє найбільші значення всіх основних характеристик геополітичного ризику. Середній рівень індексу у цьому кластері становить ($G\bar{P}R_i$), а максимальне значення досягає ($\max(GPR)_i$). Волатильність ризику є значно вищою порівняно

з іншими групами ($\sigma(GPR)_i$), що свідчить про різкі та часті зміни геополітичної ситуації. Величина шоку 2022 року є максимальною (Δ_i), що відображає сильний вплив масштабних геополітичних подій на країни цієї групи. Крім того, позитивний тренд (β_i) та значне зростання середньої річної зміни індексу (m_i) свідчать про загальну тенденцію до посилення геополітичної напруженості у досліджуваний період. Це означає, що для цих країн геополітичні конфлікти є системним фактором, який суттєво впливає на їхню політичну та економічну стабільність.

Розподіл країн за сформованими групами ризику наведено (табл. 4).

Таблиця 4 – Класифікація країн за рівнем ризику геополітичних змін методом k-means
Table 4 – Classification of countries by the level of risk of geopolitical changes using the k-means method

Країна	Група ризику	\bar{GPR}_i	$\max(GPR)_i$	$\sigma(GPR)_i$	β_i	Δ_i	m_i
Туніс	Дуже низький	0,0137	0,0241	0,0049	0,0011	0,0034	0,0001
Португалія	Дуже низький	0,0172	0,0258	0,0049	0,0017	0,0130	0,0009
Чилі	Дуже низький	0,0172	0,0278	0,0080	-0,0005	0,0125	-0,0026
Малайзія	Дуже низький	0,0262	0,0301	0,0025	0,0009	0,0055	0,0012
Аргентина	Дуже низький	0,0295	0,0412	0,0114	0,0021	0,0167	0,0005
Перу	Дуже низький	0,0309	0,0471	0,0137	0,0025	0,0254	-0,0017
Таїланд	Дуже низький	0,0354	0,0471	0,0081	0,0008	-0,0063	-0,0007
Колумбія	Дуже низький	0,0379	0,0775	0,0193	-0,0052	-0,0195	-0,0074
Філіппіни	Дуже низький	0,0402	0,0589	0,0109	0,0033	-0,0037	0,0020
В'єтнам	Дуже низький	0,0582	0,0772	0,0164	0,0054	0,0164	0,0010
Бразилія	Дуже низький	0,0796	0,1055	0,0265	0,0061	0,0291	-0,0006
Індонезія	Низький	0,0519	0,0935	0,0270	0,0094	0,0657	0,0036
Данія	Низький	0,0537	0,0940	0,0241	0,0090	0,0616	0,0044
Швейцарія	Низький	0,0635	0,0928	0,0183	0,0076	0,0456	0,0070
Південна Африка	Низький	0,0667	0,1289	0,0351	0,0196	0,0197	0,0200
Угорщина	Низький	0,0831	0,2081	0,0637	0,0212	0,1768	0,0146
Фінляндія	Низький	0,0834	0,2336	0,0739	0,0202	0,2068	0,0102
Норвегія	Низький	0,0922	0,1617	0,0372	0,0097	0,0977	0,0020
Іспанія	Низький	0,1046	0,1694	0,0346	0,0085	0,0914	0,0047
Швеція	Низький	0,1078	0,2342	0,0652	0,0197	0,1793	0,0096
Нідерланди	Низький	0,1089	0,1789	0,0378	0,0133	0,1029	0,0087
Австралія	Низький	0,1641	0,2375	0,0376	0,0102	0,0963	0,0076
Італія	Низький	0,1887	0,3503	0,0766	0,0174	0,2130	0,0138
Індія	Низький	0,2349	0,3231	0,0497	0,0108	0,1224	-0,0016
Венесуела	Середній	0,0869	0,2003	0,0526	-0,0201	-0,0357	-0,0254
Мексика	Середній	0,0957	0,1440	0,0274	-0,0021	-0,0137	-0,0054
Гонконг	Середній	0,1402	0,2082	0,0415	-0,0169	-0,0024	-0,0130
Саудівська Аравія	Середній	0,2460	0,3969	0,0870	-0,0059	-0,0195	-0,0252
Японія	Середній	0,2573	0,3833	0,0699	-0,0008	0,1538	-0,0196
Південна Корея	Середній	0,3069	0,3959	0,0695	0,0047	0,0870	-0,0115
Туреччина	Середній	0,3165	0,4452	0,0773	0,0054	0,1716	-0,0077
Польща	Високий	0,2237	0,6737	0,2087	0,0419	0,5789	0,0187
Бельгія	Високий	0,2417	0,4924	0,1149	0,0174	0,3106	0,0065
Єгипет	Високий	0,2421	0,5620	0,1792	0,0896	0,0226	0,0886
Канада	Високий	0,2574	0,3994	0,0789	0,0256	0,2100	0,0122
Франція	Високий	0,6004	0,9206	0,1627	0,0517	0,4488	0,0381
Німеччина	Високий	0,6209	1,2531	0,2928	0,0593	0,8240	0,0257
Ізраїль	Високий	0,6397	1,9173	0,6375	0,3093	0,0165	0,3355
Китай	Високий	0,9090	1,1764	0,1584	0,0394	0,3924	0,0015
Велика Британія	Високий	1,1766	1,7720	0,3714	0,1467	0,9261	0,1086
Україна	Високий	1,2295	3,6086	1,2345	0,4137	3,4043	0,2551
Країна-агресор	Високий	1,5723	3,6904	1,0575	0,3112	2,9509	0,1642
Сполучені Штати	Високий	2,5642	3,5319	0,6591	0,2629	1,5706	0,1707

Джерело: розраховано автором / Source: calculated by the author

До кластеру високого геополітичного ризику віднесено країни з найвищими значеннями індексу GPR, значною волатильністю та сильною реакцією на глобальні геополітичні події. Україна та країна-агресор демонструють найвищі значення шоку 2022 року, що пов'язано з повномасштабною війною та різкою зміною безпекового середовища в Європі. Для України характерні максимальні значення індексу та висока волатильність, тоді як у країні-агресорі високий ризик також зумовлений санкційним тиском, міжнародними політичними подіями та активною присутністю в глобальному інформаційному просторі. До цієї групи належать також США, Велика Британія, Китай, Ізраїль, Франція, Німеччина та інші ключові гравці світової політики. Високі значення GPR для них значною мірою пояснюються методологією індексу, що базується на аналізі міжнародних новин: активна участь у політичних, військових та дипломатичних процесах підвищує частоту згадувань і, відповідно, рівень індексу.

Кластер середнього ризику включає країни з помірними значеннями індексу та відносно стабільною, але не повністю передбачуваною ситуацією. До нього належать Венесуела, Мексика, Гонконг, Саудівська Аравія, Японія, Південна Корея та Туреччина. Для цих держав характерні періодичні напруження, регіональні конфлікти або політичні трансформації, що спричиняють короткострокові зростання ризику, однак без системного характеру, притаманного високоризиковому кластеру.

Кластер низького ризику об'єднує країни зі стабільним геополітичним середовищем і відсутністю значних конфліктних потрясінь. До нього належать Данія, Швейцарія, Фінляндія, Норвегія, Іспанія, Швеція, Нідерланди, Австралія, Італія та Індія. Хоча ці країни можуть брати участь у міжнародних процесах, рівень геополітичної напруженості для них залишається низьким, а зміни індексу GPR не є різкими чи системними.

Кластер дуже низького ризику включає країни з мінімальними значеннями індексу та високою стабільністю геополітичного середовища. До нього належать Туніс, Португалія, Чилі, Малайзія, Аргентина, Перу, Таїланд, Колумбія, Філіппіни, В'єтнам та Бразилія. Для цих держав характерна відсутність масштабних міжнародних конфліктів і геополітичних потрясінь у досліджуваній період, що відображається у низьких значеннях GPR та незначній волатильності.

Отримана класифікація підтверджує наявність суттєвої диференціації країн не лише за рівнем геополітичного ризику, але й за його структурною динамікою. Виявлені кластери відображають багатовимірний характер ризику, який формується як результат поєднання рівневих характеристик (GPR_i) екстремальних проявів ($\max(GPR)_i$), нестабільності ($\sigma(GPR)_i$), трендової складової (β_i) та реакції на глобальні шоки (Δ_i). Таким чином, кластеризація дозволяє перейти від статичного вимірювання ризику до його більш комплексної структурної оцінки.

З економічної точки зору це означає, що країни різних кластерів функціонують у принципово різних умовах невизначеності. Високий рівень геополітичного ризику може трансформуватися у зростання премій за ризик, підвищення вартості запозичень, зниження інвестиційної активності та уповільнення економічного розвитку. Натомість країни з низьким або дуже низьким рівнем ризику мають більш передбачуване інституційне та політичне середовище, що створює сприятливі умови для довгострокових інвестицій та економічного зростання.

Отже, результати кластеризації логічно доповнюють прогностичні оцінки. Якщо до цього було здійснено кількісне прогнозування індексу геополітичного ризику, то в межах даної кластеризації сформовано структурну класифікацію країн за рівнем та характером геополітичної напруженості. У сукупності це дозволяє отримати комплексне уявлення про макроекономічну природу геополітичних ризиків у сучасних умовах глобальної нестабільності.

Висновки. У статті досліджено можливості застосування методів машинного навчання для макроекономічного прогнозування геополітичних змін на основі панельних макроекономічних даних. У результаті проведеного дослідження розроблено та реалізовано підхід до прогнозування індексу геополітичного ризику на основі широкого набору макроекономічних показників із використанням ансамблевих алгоритмів машинного навчання – XGBoost та LightGBM. Застосування зазначених моделей дозволило врахувати складну структуру взаємозв'язків між змінними, зокрема їх нелінійний характер, що є суттєвою перевагою порівняно з традиційними економетричними підходами.

Емпіричні результати свідчать про високу точність прогнозування та підтверджують висунуту у дослідженні гіпотезу щодо ефективності використання ансамблевих алгоритмів машинного навчання для аналізу геополітичних процесів. При цьому найвищу точність прогнозування продемонструвала конфігурація моделі XGBoost на основі незаповненої бази даних, яка забезпечила найменше значення показника MAPE серед усіх розглянутих варіантів.

Важливим результатом є підтвердження здатності моделей машинного навчання ефективно працювати з високимірними даними та забезпечувати надійні прогнози навіть за наявності пропущених значень. Це розширює можливості їх застосування у задачах макроекономічного аналізу, де інформаційні бази часто є неповними або неоднорідними.

Додатково показано, що методи кластерного аналізу дозволяють доповнити результати прогнозування шляхом структуризації країн за рівнем геополітичного ризику, що сприяє глибшому розумінню глобальних процесів та ідентифікації країн із підвищеним рівнем ризику.

Отримані результати підтверджують висунуті у дослідженні гіпотези. Зокрема, встановлено, що ансамблеві алгоритми машинного навчання здатні забезпечувати високу точність прогнозування індексу геополітичного ризику та ефективно враховувати складні нелінійні взаємозв'язки між макроекономічними показниками. Також підтверджено, що використання високовимірних панельних даних сприяє підвищенню якості прогнозування завдяки можливості врахування значної кількості макроекономічних факторів та їх взаємодії.

Разом із тим, дослідження має певні обмеження. Зокрема, у роботі використано фіксований набір гіперпараметрів моделей та базовий хронологічний підхід до валідації прогнозів, тоді як застосування більш складних процедур оптимізації та тестування моделей потенційно може підвищити точність результатів. Крім того, дослідження базується переважно на макроекономічних показниках і не враховує повною мірою текстові, поведінкові чи фінансові джерела інформації, які також можуть впливати на динаміку геополітичного ризику.

Перспективи подальших досліджень полягають у розширенні набору використаних даних, зокрема за рахунок інтеграції текстових джерел інформації, новинних потоків та індикаторів економічної невизначеності. Доцільним є також застосування більш складних моделей машинного навчання, зокрема методів глибокого навчання, а також використання розширених процедур валідації прогнозів. Перспективним напрямом є інтеграція моделей прогнозування геополітичного ризику у системи підтримки прийняття рішень у сфері державної економічної політики та макроекономічного прогнозування.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. International Monetary Fund. World Economic Outlook. 2025. URL: <https://www.imf.org/en/publications/weo>
2. Міжнародні економічні відносини: Навчальний посібник / А.І. Боярчук, Р.П. Огородник, І.А. Плющик, Н.М. Антофій, Н.Є. Федорова. Херсон, ТОВ «ВКФ «СТАР» ЛТД», 2018. 374 с.
3. Gupta R., Gozgor G., Kaya H., Demir E. Effects of geopolitical risks on trade flows: evidence from the gravity model. Eurasian Econ Rev. 2019. Vol. 9. P. 515–530. <https://doi.org/10.1007/s40822-018-0118-0>
4. World Bank Group. Worldwide Governance Indicators. 2023. URL: <https://www.worldbank.org/en/publication/worldwide-governance-indicators>
5. Geopolitical Risk (GPR) Index. 2018. URL: <https://www.matteoiacoviello.com/gpr.htm>
6. World Bank Group. Global Economic Prospects. 2025. URL: <https://openknowledge.worldbank.org/server/api/core/bitstreams/0e685254-776a-40cf-b0ac-f329dd182e9b/content>
7. United Nations. World Economic Situation and Prospects. 2025. URL: <https://desapublications.un.org/publications/world-economic-situation-and-prospects-2025>
8. United Nations. Transforming our world: the 2030 Agenda for Sustainable Development. 2015. URL: <https://sdgs.un.org/2030agenda>
9. Caldara D., Conlisk S., Iacoviello M., Penn M. Do Geopolitical Risks Raise or Lower Inflation? Journal of International Economics. 2025. Vol. 159. <https://doi.org/10.1016/j.jinteco.2025.104188>

10. Hossain A. T., Masum A.-A., Saadi S. The impact of geopolitical risks on foreign exchange markets: Evidence from the Russia–Ukraine war. *Finance Research Letters*. 2024. Vol. 59, 104750. <https://doi.org/10.1016/j.frl.2023.104750>
11. Bouri E., Gabauer D., Gupta R., Kinatader H. Geopolitical Risk and Inflation Spillovers across European and North American Economies. 2023. URL: <https://repository.up.ac.za/server/api/core/bitstreams/016aa5df-47c7-4df0-9203-ab1da302b9dd/content>
12. International Monetary Fund. Global Financial Stability Report, 2025. URL: <https://www.imf.org/en/publications/gfsr>
13. Fund for Peace. Fragile States Index. 2023. URL: <https://fragilestatesindex.org/>
14. Caldara D., Iacoviello M. Measuring Geopolitical Risk. *American Economic Review*. 2022. vol. 112, no. 4. P. 1194–1225. <https://doi.org/10.1257/aer.20191823>
15. World Bank Group. World Development Indicators. 2024. URL: <https://databank.worldbank.org/source/world-development-indicators>
16. Mullainathan, S., Spiess, J. Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*. 2017. vol. 31, no. 2. P. 87–106. <https://doi.org/10.1257/jep.31.2.87>
17. Goulet Coulombe P., Leroux M., Stevanovic D., Surprenant S. How is Machine Learning Useful for Macroeconomic Forecasting? 2020. URL: <https://arxiv.org/pdf/2008.12477>
18. Xue Y., Schincariol T., Chadeaux T., Groen D. Using machine learning to forecast conflict events for use in forced migration models. 2025. *Sci Rep*. Vol. 15, 28202. <https://doi.org/10.1038/s41598-025-11812-2>
19. Athey S., Imbens G.W. Machine Learning Methods Economists Should Know About. *Annual Review Economics*. 2019. Vol. 11. P. 685-725. <https://doi.org/10.1146/annurev-economics-080217-053433>
20. Tenorio J., Perez W. Monthly GDP nowcasting with Machine Learning and Unstructured Data. 2024. URL: <https://arxiv.org/pdf/2402.04165>
21. Al-Karkhi M.I., Rządkowski G. Innovative machine learning approaches for complexity in economic forecasting and SME growth: A comprehensive review. *Journal of Economy and Technology*. 2025. Vol. 3. P. 109-122. <https://doi.org/10.1016/j.ject.2025.01.001>
22. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016 (p. 785–794). <https://doi.org/10.1145/2939672.2939785>
23. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.-L. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
24. Stryker C., Bergmann D., Lee F., Scapicchio M. IBM. The 2026 Guide to Machine Learning. 2025. URL: <https://www.ibm.com/think/topics/gradient-boosting>
25. XGBoost Documentation. 2025. URL: <https://xgboost.readthedocs.io/en/stable/tutorials/model.html>
26. LightGBM Documentation. 2025. URL: <https://lightgbm.readthedocs.io/en/latest/Features.html>
27. Ortiz A., Rodrigo T. Geopolitics, Geoeconomics and Risk: A Machine Learning Approach. 2025. URL: <https://arxiv.org/pdf/2510.12416v1>
28. Gupta R., Karmakar S., Pierdzioch, C. Safe Havens, Machine Learning, and the Sources of Geopolitical Risk: A Forecasting Analysis Using Over a Century of Data. 2023. URL: https://www.econstor.eu/bitstream/10419/317932/1/10614_2023_Article_10452.pdf
29. Zhang H., Wang W., Niu Z. Geopolitical risks and crude oil futures volatility: Evidence from machine learning. *Resources Policy*. 2024. Vol. 98. <https://doi.org/10.1016/j.resourpol.2024.105374>
30. OECD Data Portal. 2024. URL: <https://www.oecd.org/en/data.html>
31. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning. 2021. URL: https://www.karlin.mff.cuni.cz/~pesta/NMFM334/StatLearning/Book2nd/ISLRv2_website.pdf
32. Huang Z., Liang Z., Zhou S., Zhang S. An Improved Density-Based Spatial Clustering of Applications with Noise Algorithm with an Adaptive Parameter Based on the Sparrow Search Algorithm. 2025. Vol. 18(5), 273. <https://doi.org/10.3390/a18050273>

33. Hyndman, R.J., Athanasopoulos, G. Forecasting: Principles and Practice. 2021. URL: <https://otexts.com/fpp3/>

34. Rousseeuw P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. 1987. Vol. 20. P. 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)

Конфлікт інтересів: автори засвідчують, що, незважаючи на те, що один із авторів статті є член редакційної колегії цього журналу, процес рецензування, прийняття рішення щодо публікації та редагування проводилися незалежно, без їх участі чи впливу. Рецензування, остаточне рішення ухвалювалося іншими членами редакційної колегії, які не є співавторами. Будь-які потенційні конфлікти інтересів були повністю усунені шляхом зовнішнього контролю процесу.

Стаття надійшла до редакції 01.03.2026
Стаття рекомендована до друку 20.04.2026
Стаття опублікована 25.05.2026

REFERENCES

1. International Monetary Fund. (2025). World Economic Outlook. Retrieved from <https://www.imf.org/en/publications/weo>
2. Boiarchuk, A. I., Ohorodnyk, R. P., Pliushchuk, I. A., Antofii, N. M., & Fedorova, N. Ye. (2018). International economic relations. Kherson: VKF "STAR" LTD. (in Ukrainian)
3. Gupta, R., Gozgor, G., Kaya, H., & Demir, E. (2019). Effects of geopolitical risks on trade flows: evidence from the gravity model. *Eurasian Econ Rev* 9, 515–530. <https://doi.org/10.1007/s40822-018-0118-0>
4. World Bank Group. (2023). Worldwide Governance Indicators. Retrieved from <https://www.worldbank.org/en/publication/worldwide-governance-indicators>
5. Caldara, D., & Iacoviello, M. (2018). Geopolitical Risk (GPR) Index. Retrieved from <https://www.matteoiacoviello.com/gpr.htm>
6. World Bank Group. (2025). Global Economic Prospects. Retrieved from <https://openknowledge.worldbank.org/server/api/core/bitstreams/0e685254-776a-40cf-b0ac-f329dd182e9b/content>
7. United Nations. (2025). World Economic Situation and Prospects. Retrieved from <https://desapublications.un.org/publications/world-economic-situation-and-prospects-2025>
8. United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. Retrieved from <https://sdgs.un.org/2030agenda>
9. Caldara, D., Conlisk, S., Iacoviello, M., & Penn, M. (2025). Do geopolitical risks raise or lower inflation? *Journal of International Economics*, 159. <https://doi.org/10.1016/j.jinteco.2025.104188>
10. Hossain, A. T., Masum, A.-A., & Saadi, S. (2024). The impact of geopolitical risks on foreign exchange markets: Evidence from the Russia-Ukraine war. *Finance Research Letters*, 59, 104750. <https://doi.org/10.1016/j.frl.2023.104750>
11. Bouri, E., Gabauer, D., Gupta, R., & Kinateder, H. (2023). Geopolitical risk and inflation spillovers across European and North American economies. Retrieved from <https://repository.up.ac.za/server/api/core/bitstreams/016aa5df-47c7-4df0-9203-ab1da302b9dd/content>
12. International Monetary Fund. (2025). Global Financial Stability Report. Retrieved from <https://www.imf.org/en/publications/gfsr>
13. Fund for Peace. (2023). Fragile States Index. Retrieved from <https://fragilestatesindex.org/>
14. Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4), 1194–1225. <https://doi.org/10.1257/aer.20191823>
15. World Bank Group. (2024). World Development Indicators. Retrieved from <https://databank.worldbank.org/source/world-development-indicators>
16. Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106. <https://doi.org/10.1257/jep.31.2.87>
17. Goulet Coulombe, P., Leroux, M., Stevanovic, D., & Surprenant, S. (2020). How is machine learning useful for macroeconomic forecasting? Retrieved from <https://arxiv.org/pdf/2008.12477>

18. Xue, Y., Schincariol, T., Chadefaux, T., & Groen, D. (2025). Using machine learning to forecast conflict events for use in forced migration models. *Sci Rep*, 15, 28202. <https://doi.org/10.1038/s41598-025-11812-2>
19. Athey, S., & Imbens, G. W. (2019). Machine learning methods economists should know about. *Annual Review Economics*, 11, 685-725. <https://doi.org/10.1146/annurev-economics-080217-053433>
20. Tenorio, J., & Perez, W. (2024). Monthly GDP nowcasting with machine learning and unstructured data. Retrieved from <https://arxiv.org/pdf/2402.04165>
21. Al-Karkhi, M. I., & Rządowski, G. (2025). Innovative machine learning approaches for complexity in economic forecasting and SME growth: A comprehensive review. *Journal of Economy and Technology*, 3, 109-122. <https://doi.org/10.1016/j.ject.2025.01.001>
22. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. (p. 785–794) <https://doi.org/10.1145/2939672.2939785>
23. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-L. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
24. Stryker, C., Bergmann, D., Lee, F., & Scapicchio, M. (2025). The 2026 guide to machine learning. IBM. Retrieved from <https://www.ibm.com/think/topics/gradient-boosting>
25. XGBoost Documentation. (2025). Retrieved from <https://xgboost.readthedocs.io/en/stable/tutorials/model.html>
26. LightGBM Documentation. (2025). Retrieved from <https://lightgbm.readthedocs.io/en/latest/Features.html>
27. Ortiz, A., & Rodrigo, T. (2025). Geopolitics, geoeconomics and risk: A machine learning approach. Retrieved from <https://arxiv.org/pdf/2510.12416v1>
28. Gupta, R., Karmakar, S., & Pierdzioch, C. (2023). Safe havens, machine learning, and the sources of geopolitical risk: A forecasting analysis using over a century of data. *Computational Economics*. Retrieved from https://www.econstor.eu/bitstream/10419/317932/1/10614_2023_Article_10452.pdf
29. Zhang, H., Wang, W., & Niu, Z. (2024). Geopolitical risks and crude oil futures volatility: Evidence from machine learning. *Resources Policy*, 98. <https://doi.org/10.1016/j.resourpol.2024.105374>
30. OECD. (2024). OECD Data Portal. Retrieved from <https://www.oecd.org/en/data.html>
31. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An introduction to statistical learning (2nd ed.). Retrieved from https://www.karlin.mff.cuni.cz/~pesta/NMFM334/StatLearning/Book2nd/ISLRv2_website.pdf
32. Huang, Z., Liang, Z., Zhou, S., & Zhang, S. (2025). An improved density-based spatial clustering of applications with noise algorithm with an adaptive parameter based on the sparrow search algorithm. *Algorithms*, 18(5), 273. <https://doi.org/10.3390/a18050273>
33. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). Retrieved from <https://otexts.com/fpp3/>
34. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)

Conflict of Interest: The authors certify that although one of the authors of the article is a member of the editorial board of this journal, the peer-review process, the publication decision, and the editing were conducted independently, without their participation or influence. The peer review and the final decision were carried out by other members of the editorial board who are not co-authors. Any potential conflicts of interest were fully mitigated through external oversight of the process.

*The article was received by the editors 01.03.2026
The article is recommended for printing 20.04.2026
The article was published on 25.05.2026*

A. STAVYTSKYI*, D.Sc. (Economics), Professor, Professor of the Department of Economic Cybernetics, <https://orcid.org/0000-0002-5645-6758>, a.stavytskyi@gmail.com
O. BABKOVA*, Student of the Department of Economic Cybernetics, <https://orcid.org/0009-0002-3254-0565>, olha.babkova@gmail.com

* Taras Shevchenko National University of Kyiv, 60 Volodymyrska Street, Kyiv, 01033, Ukraine

MACROECONOMIC FORECASTING OF GEOPOLITICAL CHANGES IN THE WORLD USING MACHINE LEARNING METHODS

The article explores the possibilities of using machine learning methods for macroeconomic forecasting of geopolitical changes based on panel macroeconomic data. The relevance of the study is due to the increasing level of geopolitical instability in the world economy, the intensification of international conflicts, sanctions restrictions and structural crises, which significantly affect the economic development of countries. In such conditions, the need for the use of modern analytical approaches increases, capable of providing more accurate forecasting of geopolitical risks and identifying complex nonlinear relationships between macroeconomic indicators. The purpose of the study is to develop and implement an approach to forecasting the geopolitical risk index using machine learning methods. The work uses ensemble gradient boosting algorithms XGBoost and LightGBM, which allow for effective work with high-dimensional data and take into account the complex structure of relationships between variables. The MAE, RMSE and MAPE metrics were used to assess the quality of forecasting. The information base of the study was formed in the form of a panel structure "country-year", covering 43 countries for the period 2000-2024 and including 499 macroeconomic indicators that characterize production, consumption, foreign trade, institutional development, demographic and environmental processes. The target variable is the geopolitical risk index. To ensure the correctness of the forecast, the sample was divided into training (2000-2018) and test (2019-2024) periods. The results obtained indicate high forecasting accuracy for both models, the MAPE value does not exceed 1.9%, which confirms the effectiveness of the use of ensemble machine learning algorithms in the tasks of forecasting geopolitical processes. It was established that the used models demonstrate the stability of the results even in the presence of missing values and structural breaks in the time series. Additionally, cluster analysis was applied, which allowed us to identify groups of countries with different levels of geopolitical tension. The study confirmed the feasibility of using machine learning methods to analyze and forecast geopolitical changes. The proposed approach can be used as a decision-making support tool in the field of macroeconomic forecasting and economic policy.

Keywords: **machine learning, geopolitical risk, macroeconomic forecasting, ensemble algorithms, classification of countries.**

JEL Classification: C53, C55, C63, F51, O57.

Як цитувати: Ставицький А.В., & Бабкова О.М. Макроекономічне прогнозування геополітичних змін у світі з використанням методів машинного навчання. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Економічна»*. 2026. Вип. 110. С. 5–23. <https://doi.org/10.26565/2311-2379-2026-110-01>

In cites: Stavytskyi A., & Babkova O. (2026). Macroeconomic forecasting of geopolitical changes in the world using machine learning methods. *Bulletin of V. N. Karazin Kharkiv National University. Economic Series*, (110), 5–23. <https://doi.org/10.26565/2311-2379-2026-110-01> (in Ukrainian)
