DOI:https://doi.org/10.26565/2304-6201-2023-60-03

УДК (UDC) 519.7, 004.8

Novikov Artem *PhD student, Department of Artificial Intelligence Systems Karazin Kharkiv National University, Svobody Sq 4, Kharkiv, Ukraine,61022*
email: artem.slick@gmail.com;
https://orcid.org/0009-0004-5914-7098

Smyrnov Vadym *Master's student, Department of Artificial Intelligence Systems Karazin Kharkiv National University, Svobody Sq 4, Kharkiv, Ukraine,61022*
e-mail:vadym.smyrnov@student.karazin.ua;
https://orcid.org/0009-0000-1743-8541

YanovskyVolodymyr *Doctor of physical and mathematical sciences; Professor of Department of Artificial Intelligence Systems*
"Institute for Single Crystals" of National Academy of Sciences, Nauky ave. 60, Kharkiv, Ukraine, 61001
e-mail:yanovsky@isc.kharkov.ua;
https://orcid.org/0000-0003-0461-749X

Shadow zones of an artificial neuron

The extremely widespread use of artificial neural networks in the diverse areas of application makes the study of their fundamental properties highly relevant. Such studies can be used to improve the properties of neural networks.

The key goal of the work: to determine the general properties of artificial neurons and detect the presence of zones where the field of output signals has a complex fractal structure in the space of all input signals.

Research methods: to explore the space of all input signals, a software that allows modelling the neuron's response to all possible input signals with a certain length in the given alphabet has been developed. With the help of the developed application the space of all input signals can be modulated and the field of output signals in this space is graphically determined. By using the capability of the software to change the scale of the input signal space, zones with a self-similar, fractal structure have been found.

Results: it has been established that when considering the overall arrangement of the neuron's input signal space, specific areas – shadow zones – are present, which exhibit a complex fractal structure of output signal field. The impact of modifying theneuron's weights and threshold on the presence and location of such zones has been established. The changes that follow an increase in the length of the input signals have been described. The fractal dimension of the structures within shadow zones has been determined.

Conclusions: the obtained general properties of neurons should significantly impact the properties of neural networks in the form of shadow zones in which the "response" of the network is extremely sensitive even to minute alterations in input signals. The presence of such zones is an extremely important factor that needs to be considered while developing neural networks.

Keywords: *artificial neuron, weights, fractals, fractal dimension, activation function, space of input signals, field of output signals.*

Як цитувати: Novikov A., Smyrnov V., Yanovsky V., Shadow zones of an artificial neuron . *Вісник Харківського національного університету імені В.Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2023. вип. 60. С.27-35. https://doi.org/10.26565/2304-6201-2023-60-03

How to quote: Novikov A., Smyrnov V., Yanovsky V., "Shadow zones of an artificial neuron", *Bulletin of V.N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 60, pp.27-35, 2023. https://doi.org/10.26565/2304-6201-2023-60-03 .

1. Introduction

Nowadays, neural networks have garnered significant attention as one of the promising domains in the advancement of artificial intelligence which has been introduced in the scientific study as a very simplified mathematical model of a natural neuron. It represented a nonlinear function, which is an activation function from a linear combination of input signals. In the research by Warren McCulloch and Walter Pitts [1], the Heaviside function played the role of the activation function, and such neurons could execute logical operations, and when organized into a neural network, they demonstrated the capability to conduct numerical calculations. It was planned to construct neural networks based on electronic circuits, where neurons would function using binary signals. That corresponded to a certain degree with natural neurons, which could be in an excited or non-excited state. The practical implementation of such neural network was achieved by Frank Rosenblatt in the form of a perceptron [2, 3]. Rosenblatt anticipated substantial success in the advancement of artificial intelligence through neural networks.

The development of neural networks demanded the creation of more versatile neurons capable of processing numerical signals. Such neurons were proposed by Withrow and Hoff [4]. They suggested using a logistic dependence, or a sigmoid function, as an activation function. An important stage in the development of neural networks was the inception of the idea of learning neural networks, which in the form of the first learning algorithm appeared in the work of Hebb [5].

However, the M. Minsky and S. Papert [6] demonstrated the limitations of single-layer perceptrons and notably reduced attention to neural networks. The interest revived after the publication of Hopfield's works [7, 8, 9], and after the discovery and advancement of the backpropagation [10] the rapid development of learning neural networks began.

Now, the study of neural networks is becoming even more important, as we are facing an ever-increasing amount of data and data-based tasks that require complex analytical and decision-making solutions. Neural networks are used in many areas, including medicine, finance, automation, image recognition, natural language, and many others [11, 12, 13]. They have become a fundamental tool for achieving and enhancing solutions that were previously performed exclusively by humans [14].

This work researches the output field of a neuron for all possible input words of a certain length. It is demonstrated that within the input word space, there are zones – shadow zones – where the output field exhibits a fractal structure. The range of scales at which a self-similar fractal structure is observed is determined by the length of the input words n, encompassing a vast scale range $1 \div 2^{-n}$. The fractal dimension of such structures is determined. An area where shadow zones emerge within neuron weight values is pinpointed. The changes that might occur as the number of input channels of the neuron increases are discussed.

2. Formulation of the problem

Let us consider how a typical neuron responds to various input words within a certain alphabet. For the sake of simplicity, we will use a binary alphabet $A=[0,1]$. Using this alphabet, we will construct input words and generate the neuron's response or output words. The main goal of research is to determine the structure of the neuron's responses to all input words.

To answer this question, let us discuss the set of all questions and our conceptualization of them. Certainly, all questions consist of words from the alphabet A, and this set is denoted as A^*. According to Cantor's theorem, the cardinality of the set A^* is the continuum. To distinguish finite subsets, we also consider sets of words of finite length n, and for such subsets, we use the notation A_n^*. Therefore, any word $x \in A_n^*$ has a length $|x|=n$. The number of elements in A_n^* is equal to 2^n. As a result, all questions can be arranged in lexicographic order, thereby forming a completely ordered set. It is easy to see that each of these words corresponds to a particular interval of length $\frac{1}{n}$ within the unit interval (see Fig. 2.1). Within limit $n \to \infty$ each infinite word corresponds to a point in the unit interval. Thus, it can be deduced that each input word encodes a number less than one in the binary numeral system. For example, the number 0.011000111101 corresponds to the input word 011000111101. It is evident that the intervals of length $\frac{1}{n}$ correspond to numbers that have the same first n signs and differ in subsequent signs. This implies that such an interval comprises truncated numbers up to the decimal point. By agreement, if the number contains fewer signs, zeroes are appended to the record to match the agreed-upon length. Such coding method allows the complete set of questions presented as an input of a neuron to be displayed graphically.

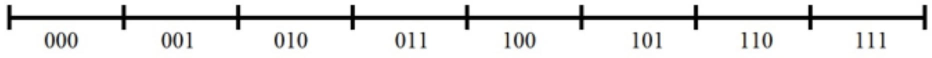

Fig. 2.1. An example of dividing a unit segment into 2^3 subintervals of length $1/2^3$ is shown to determine a mutually unique correspondence with all words of length 3 in the alphabet $A=[0,1]$. Words of any length can be arranged in the same way.

For simplicity's sake, we will consider a neuron with two input channels (see Fig. 2.2). Then each pair of input words that are sent to the input of the neuron corresponds to the cells of the unit square with coordinates (x, y). If the length of the input word is n, then the associated cell's side length is $\frac{1}{n}$. In fact, this square divided into cells, represents the space of all input words with a length of n. That enables the visualization of the neuron's response to every possible input word. To do this, let us assign a color to each cell based on the output value corresponding to the input words. To generate source words, the sequences in the binary alphabet are sent to the input channels. At the output, we will receive a sequence of the same length. There are two ways to achieve that. The first is to use the McCulloch-Pitts neuron [1], [16] with an activation function that corresponds to the Heaviside θ-function. The second is to use a standard neuron with a sigmoidal activation function, but set the rule that a value exceeding $1/2$ corresponds to the reaction 1, and if it is less than $1/2$, then it corresponds to the reaction 0. In both cases, the neuron's output will consist of a sequence of binary symbols with the same length as the input words.

Thus, each pair of input words relates to a cell of the size $\frac{1}{|x|}$ which is divided by the unit square (see Fig. 2.2). From a number coding perspective, such a cell corresponds to all real numbers that have the same first $|x|$ decimal places and differ in subsequent signs. In order to determine the neuron's field of reactions to all possible input words, it is enough to color each cell according to the initial sequence. Thus, a particular palette should be used. In our research, we have used the simplest palette. So, if the initial sequence is encoded with a number less than $1/2$, then the response corresponds to yellow, otherwise, it corresponds to blue. Next, we have used a program that executes such an algorithm automatically. By providing all possible words to the input of the neuron, it determines the output word and sets the color of the corresponding cell within the unit square. After processing all input words, it depicts a single painted square, i.e. the output field. A more detailed description of the program is provided in Appendix A. In short, its functionality corresponds to sending all possible pairs of input words to the neuron's input, obtaining the output sequence and coloring the corresponding cell in a certain color based on the resulting output word. As a result, we get the response field of the neuron.

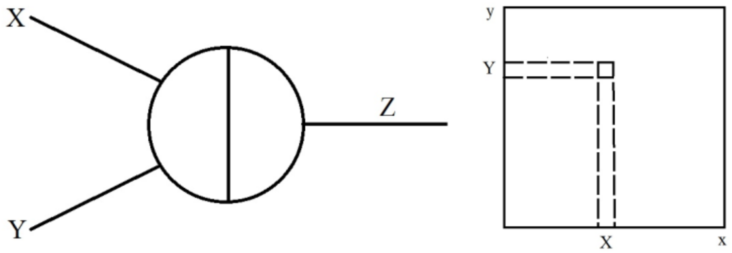

Fig 2.2. An artificial neuron with two input channels is symbolically shown on the left side. X and Y inputs are binary sequences of the same length, Z output is a binary sequence of the same length. On the right side, the positioning of the two input sequences X and Y on the unit square is shown.

3. Modeling results

Using the developed program (see Appendix A), the response field of an artificial neuron with a sigmoid activation function and a McCulloch-Pitts neuron with an activation function that corresponds to the Heaviside function at certain weight values can be examined. The input words had a length of 8. The resulting response field is shown in Fig. 3.1. It is easy to see the extraordinary structure of this field for both neurons. The input word space is divided into three characteristic zones, each demonstrating a unique structure of output words. The blue zone has no structures and occupies 3/4 of the unit square's area. The triangular yellow zone, lacking distintive structures, encompasses 1/8 of the unit square's areas. The final zone that occupies 1/8 of the area is structured and corresponds to the fractal structure of the response field. This class of objects is extremely common in nature and has been the subject of exensive study since the work of Mandelbrot [17]. Further, the zones in the space of input words in which the output words form fractal structures are referred to as "shadow zones". The fractal structure observed in these zones is well known as the Sierpiński carpet [17, 18]. Fractal, self-similar objects are characterized by a non-integer dimension, commonly known as fractal dimension. For example, the fractal dimension of the structure in Fig. 3.1 is $D_f = \dfrac{\ln 3}{\ln 2} \approx 1.59$.

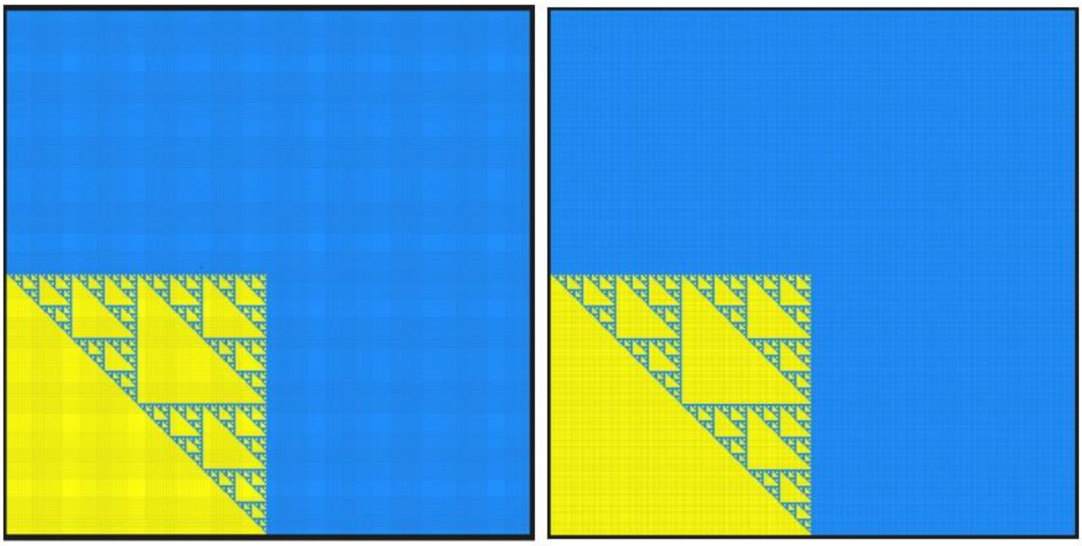

Fig 3.1. The field of responses to all possible questions with a length of 8. On the left, the field of responses of a McCulloch-Pitts neuron is depicted ($\omega_x = 1, \omega_y = 1, \sigma = -0.5$), and a regular neuron with a sigmoid activation function ($\omega_x = 1, \omega_y = 1, \sigma = -0.2$) is presented on the right.

Thus, the response of a regular neuron to all possible input words has shadow zones with fractal structures, and that has extremely interesting consequences. Indeed, if the questions are located in the non-fractal zones, the neuron's response remains unchanged with small alterations in the input words. Conversely, in the shadow zones, even a minor modification to the last character of the input word can result in a drastic change of the neuron's response. For instance, assuming that blue indicates "Yes", and yellow indicates "No", a modification in the last character of the input word may change the response from "Yes" to "No". In other words, such sensitivity can be considered as the manifestation of chaos in the neuron's responses. This is a distinct form of chaos that does not stem from random influences on the neuron, but is its internal property and refers to deterministic chaos [19, 20].

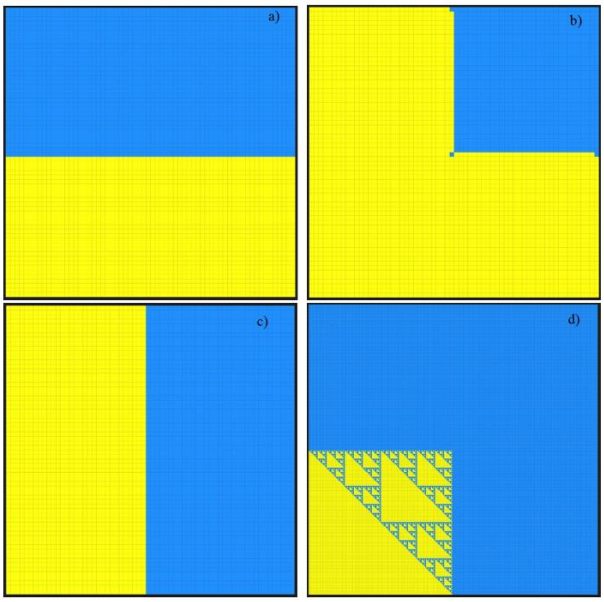

Fig. 3.2. All typical response fields of the neuron observed during simulation at different values of the neuron's weights are shown, excluding trivial cases of one color only.

Let us consider how typical the appearance of shadow zones with fractal structures is. To achieve thats, we have determined the neuron weights associated with the presence or absence of shadow zones. The result of such check can be depicted graphically. Fig. 3.2 shows all possible response fields of the neuron (excluding the trivial case where the entire space of input words is yellow*)*,denoted as *a), b), c),* and *d)* corresponding to different weight values obtained from the simulation. The first three response fields do not have shadow zones and, consequently, fractal structures, while the fourth field demonstrates a typical example with a shadow zone. It is important to note that case *b)* depends on the length of the input words. Namely, the small squares positioned near the corners of the big blue square have the size $\frac{1}{2^n}$.That is, as the length of the input words increases, the squares decrease in size and become less noticeable. No other structures have been observed in the simulation. Analyzing the square in the coorqdinate axes (ω_1, ω_2), we have determined the weights that result in the structures of the response field depicted in Fig. 3.2. The results of this analysis are shown in Fig. 3.3. The range of weights from 0 to 1 shows the weights at which fractal structures are observed. This zone is highlighted in green in Fig. 3.3. Thus, in the case of a threshold $\sigma = 0.2$, it covers 0.69 of the plane of the considered weight change. The non-fractal structures encompass 0.31 of the square area.

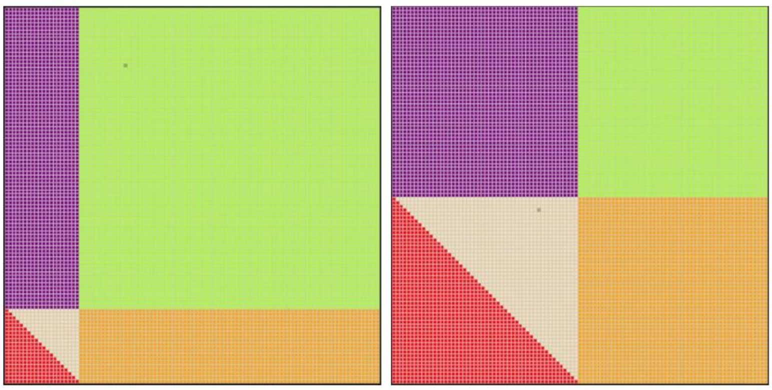

Fig. 3.3. The regions of the neuron's weights ω_1, ω_2 in which response fields have been observed. The green area represents the neuron weight values where the shadow zones appear. The purple area corresponds to the field of answers - a), orange - b), light brown - c), and red - all yellow answers.

On the left side, it shows the range of neuron weights when the threshold is set to $\sigma = 0.2$. On the right side, it illustrates the range of neuron weights with a threashold $\sigma = 0.5$. The influence of the threshold on the sizes of the corresponding squares is noticeable. There is practically no difference in choosing the sigmoid or Heaviside activation function.

It can be seen, that when the weights change within the $[0,1]$ interval, shadow zones occur twice as frequently. Altering the threshold results in changing of all zone planes. Therefore, increasing the threshold makes the size of the green zone decrease. The case when the threashold is set to $\sigma = 0.5$ is shown in Fig. 3.3 on the right side. Decreasing the threshold, conversely, expands the area of the green zone as shown in Fig. 3.3 on the left side. The symmetrical structure of the square divisions remains intact, only a shift along the diagonal of the square is observed. This allows predicting the changes occuring with alterations in the threshold. Data approximation reveals a linear relationship between the area of the green zone in the unit square of the weights and the threshold value. It can be concluded that the existence of shadow zones with a fractal structure of the response field in the input word space is a common characteristic of neurons.

4. Conclusion

Thus, it has been proven that the presence of shadow zones within the input word space, displaying output words with a fractal structure, is an internal characteristic of artificial neurons. The fractal dimension of the resulting fractal structures is equal to $D_f = \frac{\ln 3}{\ln 2} \approx 1.59$. The region of weights where shadow areas are observed in the neuron's response field is determined. The area of the weight range where shadow zones are observed depends on the threshold of the neuron. As the length of the input words increases, the fractal structure persists and extends to extremely small scales ranging from 1 to 2^{-n}. This generates zones of extreme sensitivity of the neuron's responses to even the slightest alterations in input words. In those zones, a reversal of the answer can happen even when the last character in the input word is changed.

It is evident that increasing the number of input channels of an artificial neuron does not influence the existence of shadow zones in the input word space, but visualizing these zones becomes more challenging due to the expanded dimensionality of the space of all possible input questions or words, which aligns with the number of input channels of the neuron. Consequently, visualizing shadow zones and their corresponding fractal structures becomes impossible for four or more input channels. With three input channels, this problem can still be solved. Certainly, when maintaining the weight of the third input channel and the input word constant, we can observe the aforementioned fractal structures on the response field to input words that are transmitted through two channels. Geometrically, that image corresponds to a certain intersection of a unit cube divided into corresponding colored cubes of smaller sizes, namely 2^{-n}. With a larger number of input channels, it is possible to detect such fractal structures without the visualization, by using special mathematical algorithms. Such algorithms are used in the study of strange attractors in dissipative nonlinear dynamical systems in high-dimensional spaces.

Appendix A

In our study of artificial neurons we have been using a custom-built web application designed to model and visualize the behavior of two different neuron models: SigmoidNeuron and ThresholdNeuron (McCulloch-Pitts neuron). This web application has been developed using the React library for building user interfaces and TypeScript for enhanced code quality.

The interface (see Fig. 5.1) provides separate settings fields for each neuron model. It allows adjusting such parameters as signal weights (w1, w2) and bias (threshold). Additionally, users have the opportunity to set the length of the question, thereby affecting the number of combinations of inputs for the neurons. For instance, a question of length 2 generates $(2 * 2)^2 = 16$ combinations of x1 and x2.

Fig. 5.1. The application interface.

Upon clicking the "Generate" button, the application generates a table representing all possible combinations of inputs. The square visualization component represents the table, plotting x1 values horizontally from left to right and x2 values vertically from bottom to top (see Fig. 5.2). Each cell in the table corresponds to a specific combination of x1 and x2 inputs. The application sends these inputs to the neurons, resulting in a sequence of zeros and ones as the output. If the output sequence exceeds the average value of the table, the cell is colored blue, representing a positive response ("yes"). Otherwise, the cell is colored yellow, indicating a negative response ("no").

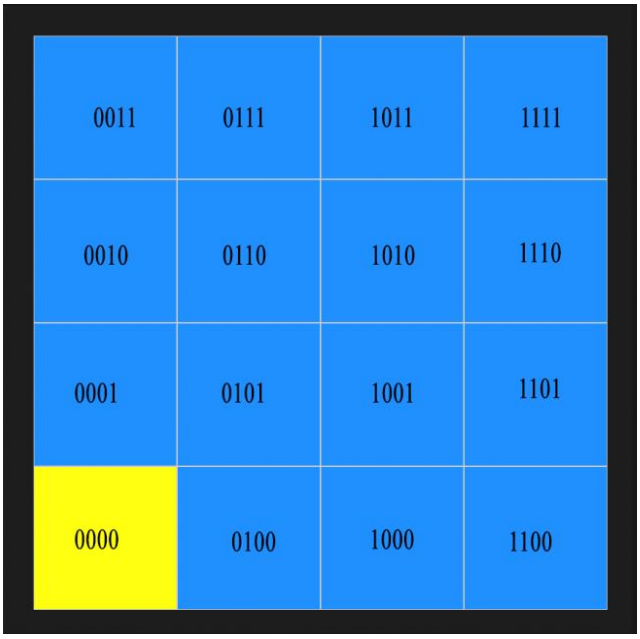

Fig. 5.2. Example of values arranging in the table for question of length 2.

Application has the following workflow:

1. User configures neuron model settings.
2. User sets the question length.
3. User clicks "Generate" to create a table of binary combinations.
4. The application processes each cell, sending inputs to the neurons and determining the output sequence.
5. Cell coloring reflects the decision outcome: blue for "yes," yellow for "no."
6. The final output is a square grid of cells, where each cell corresponds to a specific question and its color indicates the neuron's response.

This web application serves as a powerful tool for providing a visual understanding of the decision-making processes of artificial neurons, offering valuable insights into their behavior based on different configurations and input patterns.

REFERENCES

1. W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.https://doi.org/10.1007/BF02478259
2. F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958.https://doi.org/10.1037/h0042519
3. F. Rosenblatt, *Principles of Neurodynamics*. Washington, DC: Spartan Books, 1962.https://doi.org/10.2307/1419730
4. B. Widrow and M. E. Hoff, "Adaptive switching circuits," in 1960 *IRE WESTCON Conference Record*. New York, 1960.https://doi.org/10.7551/mitpress/4943.003.0012
5. D. O. Hebb, The organization of behavior; a neuropsychological theory. New York: Wiley, 1949.
6. M. L. Minsky and S. A. Papert, Perceptrons. Cambridge, MA: MIT Press, 1969.https://doi.org/10.1016/s0361-9230(99)00182-3
7. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, April 1982.https://doi.org/10.1073%2Fpnas.79.8.2554
8. J. J. Hopfield, "Neural with graded response have collective computational properties like those of two-state neurons," *Proceedings of the National Academy of Sciences*, vol. 81, no. 10, pp. 3088-3092, May 1984.https://doi.org/10.1073/pnas.81.10.3088
9. J. J. Hopfield, "Learning algorithms and probability distributions in feed-forward and feed-back networks,"*Proceedings of the National Academy of Sciences*, vol. 84, no. 23, pp. 8429-8433, December 1, 1987.
10. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing*, vol. 1, pp. 318–362. Cambridge, MA: MIT Press, 1986.https://doi.org/10.7551/mitpress/5236.003.0012
11. Y. Wu and J. Feng, "Development and application of artificial neural network," *Review of Wireless Personal Communications*, vol. 102, no. 2, pp. 1645-56, 2018.https://doi.org/10.1007/s11277-017-5224-x
12. A. Sharma and A. Chopra, "Artificial neural networks: Applications in management," *Review of Journal of Business and Management*, vol. 12, no. 5, pp. 32-40, 2013.https://doi.org/10.1371%2Fjournal.pone.0212356
13. J. Jiang, P. Trundle, and J. Ren, "Medical image analysis with artificial neural networks," *Review of Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 617-31, 2010.https://doi.org/10.1016/j.compmedimag.2010.07.003
14. A. Abraham, "Artificial Neural Networks," in *Handbook of Measuring System Design*, 1st ed., vol. 3, John Wiley & Sons, 2005. http://dx.doi.org/10.1002/0471497398.mm421
15. G. Cantor, "Über eine elementare Frage der Mannigfaltigskeitslehre," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 1, pp. 75–78, 1891.[in German]http://eudml.org/doc/144383
16. P. Halmos, *Naive Set Theory*. Princeton, NJ: D. Van Nostrand Company, 1960. Reprinted by Springer-Verlag, New York, 1974.
17. B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: W. H. Freeman and Co., 1982.
18. M. F. Barnsley and H. Rising, *Fractals Everywhere*. Boston: Academic Press Professional, 1993.
19. R. M. Crownover, *Introduction to Fractals and Chaos*. University of Missouri-Columbia: Jones and Bartlett Publishers, Boston, London, 1995.
20. H. G. Schuster, *Deterministic Chaos: An Introduction*. Weinheim: Physik-Verlag, 1984.

Зони тіні штучного нейрона

Новіков Артем Олександрович — *аспірант кафедри штучного інтелекту та програмного забезпечення Харківський національний університет імені В.Н. Каразіна; майдан Свободи, 6, м. Харків, Україна, 61022*

Смирнов Вадим Максимович — *студент магістратури кафедри штучного інтелекту та програмного забезпечення Харківський національний університет імені В.Н. Каразіна; майдан Свободи, 6, м. Харків, Україна, 61022*

Яновський Володимир Володимирович — *доктор фізико-математичних наук; професор штучного інтелекту та програмного забезпечення Інститут монокристалів, Національна Академія Наук Україн; пр. Науки 60, м. Харків, Україна, 61001*

Надзвичайна поширеність використання штучних нейронних мереж у самих різноманітних напрямках застосувань робить дослідження їх фундаментальних властивостей надзвичайно актуальним. Таки властивості можуть бути використані для покращення властивостей нейронних мереж.
Мета роботи полягає в визначенні загальних властивостей штучних нейронів та виявлені наявності у просторі всіх вхідних сигналів зон, де поле вихідних сигналів має складну фрактальну структуру.
Методи дослідження: для дослідження простору всіх вхідних сигналів було розроблено програмне забезпечення, яке дозволило моделювати реакцію нейрона на всі можливі вхідні сигнали певної довжини у деякому вхідному алфавіті. За допомогою цього забезпечення було з модульовано простір всіх вхідних сигналів та графічно визначено поле вихідних сигналів у над цим простором. Використовуючи можливість програмного забезпечення змінювати масштаб спостереження простору вхідних сигналів було здійснено пошук зон з самоподібною, фрактальною структурою.
Результати: У роботі визначено, що у випадку загального положення у просторі всіх вхідних сигналів нейрону існують зони - зони тіні, в яких поле вихідних сигналів має складну фрактальну структуру. Визначено вплив зміни ваг та порогу нейрону на існування та розташування таких зон. Виявлено зміни які спостерігаються зі збільшенням довжини вхідних сигналів. Встановлено фрактальну розмірність структур в зонах тіні.
Висновки: отримані загальні властивості нейронів повинні суттєво впливати на властивості нейронних мереж у вигляді наявності зон тіні в яких "відповідь" мережі є надзвичайно чутливою навіть до малих змін вхідних сигналів. Наявність таких зон є надзвичайно важливим фактором який потрібно враховувати при створенні неронних мереж.

Ключові слова: *штучний нейрон, ваги, фрактали, фрактальна розмірність, функція активації, простір вхідптх сигналів, поле вихідних сигналів.*

Надійшла у першій редакції 23.10.2023, в останній – 28.11.2023.
The first version has been received on 23.10.2023, the final version – on 28.11.2023.