

УДК 004.89

Лисецький Юрій Михайлович*доктор технічних наук, доцент**професор кафедри**Воєнна академія імені Євгенія Березняка**Київ, Україна**e-mail: Yurii.Lysetskyi@snt.ua**ORCID: 0000-0002-5080-1856*

ОСНОВИ ТА СКЛАДОВІ ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. Наведено, що основами та складовими штучного інтелекту є: кібернетика, теорія інформації, теорія обчислень за допомогою цифрових операцій, логіка, нейропсихологічна теорія, теорія нейронних мереж, теорія навчання заснована на кон'юнктурі нейронних мереж та синапсів, математична (комп'ютерна) модель мозку. Розглянуто нейромережі, як математичні моделі, що описують математичними формулами взаємозалежності між різними факторами і дозволяють виявити закономірності взаємозалежності вхідних даних та кінцевого результату в процесі навчання.

Ключові слова: штучний інтелект, нейромережа, модель, кібербезпека.

Вступ

Постановка проблеми. Штучний інтелект (ШІ), як галузь дослідження була заснована на семінарі, проведеному у Дартмутському коледжі в США в 1956 року [1]. Цей семінар мав важливе значення для науки: він об'єднав вчених, що цікавилися проблемою моделювання людського розуму, затвердив появу нової галузі науки і дав їй назву – artificial intelligence – штучний інтелект. Організаторами заходу були Джон Маккарті, Марвін Мінскі, Клод Шеннон і Натаніель Рочестер. Вони запросили всіх відомих американських дослідників, чия робота так чи інакше була пов'язана з питаннями теорії управління, теорії автоматів, нейронних мереж, теорії ігор і дослідженням інтелекту.

З початку 2000-х років зі стрімким розвитком інформаційних технологій ШІ починають дуже активно і широко використовувати, зокрема у інформаційній безпеці, з часом у кібербезпеці. Тому в цьому контексті висвітлення основ та складових ШІ є досить доцільним.

Аналіз останніх досліджень і публікацій. Питанням окремих теоретичних та математичних складових ШІ, присвячені наукові праці, як зарубіжних, так і вітчизняних дослідників [2–9], однак з появою можливостей використання технологій ШІ в сфері забезпечення кібербезпеки їх ефективного застосування потребує більш системного дослідження основ та складових ШІ.

Мета статті – висвітлити основи та складові ШІ.

Виклад основного матеріалу

Наука про штучний інтелект (ШІ) починає розвиватися з 40-х років минулого століття, коли група вчених з різних галузей прийшли до думки про можливість створення штучного мозку [1]. Тоді почали формуватися його теоретичні, математичні основи та складові [2]. Дослідження в галузі нейрології показали, що мозок являє собою мережу з нейронів, які обмінюються електричними сигналами за принципом “все або нічого”, 0 або 1. Кібернетика американського математика Норберта Вінера описала основи управління і стабільності в електричних мережах [3]. З теорії інформації американського

електротехніка і математика Клода Шеннона стало відомо про цифрові сигнали [4]. Англійський математик і криптограф Алан Тюринг теоретично обґрунтував можливість будь-яких обчислень за допомогою цифрових операцій [5]. Американськи нейропсихолог і нейрофізіолог Уоррен Ма-Каллок та нейrolінгвіст і математик Волтер Пітс проаналізували мережі, які складають з ідеалізованих штучних нейронів, і продемонстрували, як вони можуть виконувати найпростіші логічні функції [6]. Канадський фізіолог і нейропсихолог Дональд Гебб запропонував теорію навчання, засновану на кон'юнктурі нейронних мереж та синапсів, здатних з часом посилюватись чи слабшати [7]. Ці науковці були першими, хто описав те, що дослідники згодом назвуть нейронною мережею.

У 1957 році Френком Розенблаттом був запропонований Перцептрон (англ. perceptron) – математична або комп'ютерна модель сприйняття інформації мозком (кібернетична модель мозку), вперше втілена у вигляді електронної машини “Марк-1” у 1960 році. Перцептрон став однією з перших моделей нейромереж, а “Марк-1” – першим у світі нейрокомп'ютером [8].

Отже, нейромережа – це по суті математична модель, яка емулює роботу нервової системи людини [9]. В свою чергу, математична модель – це система математичних співвідношень, які описують досліджуваний процес або явище. Таким чином, нейромережа – це математична модель, яка описує математичними формулами взаємозалежності між різними факторами і дозволяє виявити закономірності взаємозалежності вхідних даних та кінцевого результату в процесі навчання. У математичній моделі нейромережі є вузли – штучні нейрони. Вузли з'єднані між собою зв'язками і передають сигнали один одному (рис. 1).

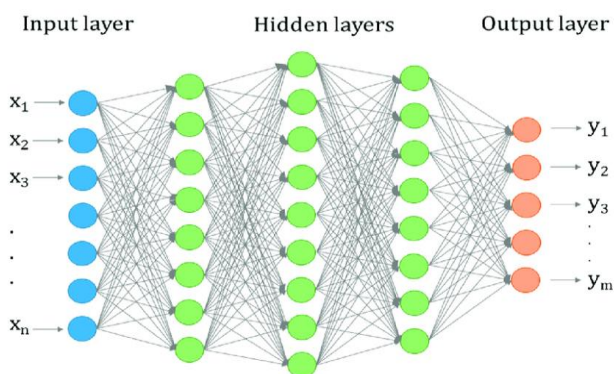


Рис. 1. Схема багатошарової нейронної мережі

Найпростіша математична модель нейрона називається Перцептрон (Perceptron) [10]. Перцептрон можна візуалізувати, як представлено на рис. 2:

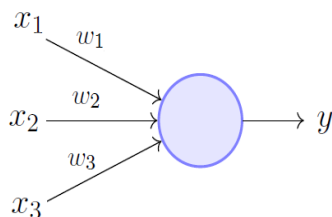


Рис. 2. Схематичне зображення перцептрону

Перцептрон – це математична функція, якій задається декілька параметрів “ x ”, а результатом обчислення є число “ y ”, наприклад $y = 3 \cdot x$. Кожному вхідному значенню “ x ” надається певна вага “ w ”, яку можна регулювати, в залежності від того, наскільки це

значення важливе та впливає на результат обчислення. В даному випадку, ця функція буде виглядати так:

$$y = \begin{cases} 0, & \text{якщо } \sum_i w_{ij} x_j \leq \text{порогове значення} \\ 1, & \text{якщо } \sum_i w_{ij} x_j > \text{порогове значення} \end{cases}$$

Насправді функція виглядає складніше: по суті вона означає, що результат роботи нейрону є 0, якщо сума добутків всіх вхідних параметрів та їх вага менше певного порогового значення або дорівнює 1, якщо сума всіх вхідних параметрів та їх вага більше певного порогового значення. 0 – нейрон пасивний, 1 – нейрон активний. Результат обчислення передається як вхідний параметри іншим нейронам наступного шару нейромережі. Чим більше “у” – тим сильніше “активується” нейрон на наступному шарі і так далі.

Отже, роботу перцептрону можна уявити як пристрій який приймає рішення на основі зважених вхідних даних. Якщо уявити, математичну модель з тисячею перцептронів, тобто тисячу нейронів у кожного на вході будуть тисячі чи десятки тисяч “х”. І чим більше вузлів, тим більшу варіативність може моделювати мережа. Керуючи вагами з’єднань між нейронами ми можемо підлаштовувати нейромережу, та таким чином навчати її видавати очікувані результати.

Перцептрон – це найпростіша, але і найменш ефективна модель вузла нейромережі. Крім перцептрону ще є лінійний нейрон, сигмоїдний нейрон, нейрон з функцією активації ReLU, нейрон з функцією активації tanh, нейрон з функцією активації softmax [11]. Кожен має специфічні функції для різних завдань нейромереж.

Архітектуру нейромережі можна поділити на дві складові: архітектуру математичної моделі нейромережі та архітектуру програмного забезпечення яке реалізує функції нейромережі. Для початка наведемо найбільш використовувані архітектури математичних моделей:

Нейромережі прямого зв'язку (Feedforward). Сигнали передаються в одному напрямку від одного шару до іншого. Це найпростіша структура призначені для класифікації чи прогнозування.

Нейромережі з зворотнім зв'язком (Recurrent). У цій структурі нейрони можуть мати зв'язки не тільки вперед, але й назад, створюючи цикли в мережі. Використовується, наприклад, у машинному перекладі, де контекст та попередній контекст важливі для правильного переклада.

Згорткові нейромережі (Convolutional). Дозволяють локальну обробку даних та розпізнавання властивостей в зображеннях. Згорткові нейромережі ефективні в розпізнаванні об'єктів та виявленні патернів в зображеннях.

Окрім вказаних, існують рекурентні нейромережі з довготривалою пам'яттю (Long Short-Term Memory, LSTM), гортково-рекурентні нейромережі (Convolutional Recurrent Neural Networks), генеративно-протилежні нейромережі (Generative Adversarial Networks), мережі асоціативної пам'яті (Hopfield Networks, Boltzmann Machines), мережі денної пам'яті (Memory Networks), генеративні адверсаріальні мережі (Generative Adversarial Networks, GAN) та інші [12]. Архітектури нейромереж постійно розвиваються та модифікуються з метою підвищення ефективності навчання, зменшення помилок в роботі, виконання складних комплексних завдань.

З точки зору прикладного програмного забезпечення, при реалізації функцій нейромережі, можна визначити три рівня:

- *Базовий рівень* – це основана частина нейромережі, яка формується в процесі основного глибокого навчання;
- *Рівень адаптації* – це адаптація під певну організацію, користувача, законодавство і цей рівень містить унікальний досвід кожної практичної реалізації нейромережі;

– *Контекстний рівень* – результат навчання нейромережі при взаємодії конкретних суб'єктів з нею, тобто контекст конкретного користувача чи організації.

Отже, нейромережі є сучасним потужним інструментом III, який здатний моделювати та симулювати роботу людського мозку. Це практичний механізм, який забезпечує реалізацію технологій Machine Learning та Deep Learning.

Machine Learning – це підгалузь III, яка вивчає алгоритми та моделі, що дозволяють комп'ютерам самостійно вчитися та вдосконалювати свою продуктивність на основі даних, без явного програмування [13].

Deep Learning – це підгалузь машинного навчання, яка використовує нейромережі зі значною кількістю шарів для автоматичного вивчення представлення даних. Глибоке навчання дозволяє комп'ютерам розуміти та аналізувати складні залежності, виявляти приховані закономірності та здійснювати передбачення на основі великого обсягу даних [14].

Різниця Deep Learning та Machine Learning – в складності завдань [15]. Deep Learning виявляє складніші залежності та використовує значно більше даних для побудови закономірностей. Компанії Google, Microsoft, Facebook впроваджують ці технології у своїх продуктах. Справжньою сенсацією стала мовна модель ChatGPT (найновіша доступна версія GPT-4), яка генерує текст, базуючись на запитах користувача та враховуючи контекст попередніх запитів [16]. Це модель глибокої нейромережі, яка базується на архітектурі GPT (Generative Pre-trained Transformer). GPT використовує трансформерну архітектуру та багатошарові рекурентні нейромережі (RNN) для генерації тексту та відповідей на запитання.

Сьогодні нейромережі застосовуються в різних технологічних продуктах та галузях, зокрема у кібербезпеці.

Висновки

Основами та складовими III є кібернетика, теорія інформації, теорія обчислень за допомогою цифрових операцій, логіка, нейропсихологічна теорія, теорія нейронних мереж, теорія навчання заснована на кон'юнктурі нейронних мереж та синапсів, математична (комп'ютерна) модель мозку. Справжній прорив у використанні III почався у 2010-х роках з розвитком машинного та глибокого навчання. Цифровізація, BigData та можливість обробляти великі масиви даних забезпечили для цього необхідний ресурс.

Отже, моделі III на основі нейромереж використовуються для розв'язання складних задач, які не завжди можна вирішити за допомогою математичних формул або правил. Вони здатні до навчання на основі великих обсягів даних та розпізнавання складних залежностей у цих даних. Можуть застосовуватись в різних технологічних продуктах та галузях, зокрема у кібербезпеці.

Список літератури

1. The History of Artificial Intelligence from the 1950s to Today. URL: <https://www.freecodecamp.org/news/the-history-of-ai/> (дата звернення 24.10.2024).
2. Лисецький Ю.М., Сокур В.В. Теоретичні та математичні основи штучного інтелекту. *Вісник воєнної розвідки*. 2024. Vol. 85. С. 51-56.
3. Norbert Wiener. *Cybernetics or Control and Communication in the Animal and the Machine*, Paris – Cambridge, Mass., 1948. 194 p.
4. Shannon C. E. A Mathematical Theory of Communication. *Bell System Technical Journal*. 1948. № 27. P. 379-423.
5. Turing A. M. Computing Machinery and Intelligence. *Mind*. 1950. № 49. P. 433-460.
6. Warren S. McCulloch and Walter Pitts, A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. 1943. Vol. 5. P. 115-133.
7. Donald Olding Hebb *The Organization of Behavior: A Neuropsychological Theory*. D.O. Hebb: The Organization of Behavior, Wiley: New York, 1949. 337 p.

8. Rosenblatt F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*. 1958. Vol. 65. № 6. P. 386-408.
9. Калбазов Д.Й., Данченко О.І., Лисецький Ю.М. Нейромережі. Розвиток та перспективи. *Математичні машини і системи*. 2024. № 2. С. 24–32.
10. What is Perceptron? A Beginners Guide for 2023. URL: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/perceptron> (дата звернення 15.11.2024).
11. 7 архитектур нейронных сетей для решения задач NLP. URL: <https://neurohive.io/ru/osnovy-data-science/7-arhitektur-nejronnyh-setej-nlp/> (дата звернення 23.09.2024).
12. Які існують типи нейронних мереж? URL: <https://devzone.org.ua/qna/iaki-isnuiut-tipi-neironnix-merez>.
13. What is machine learning? URL: <https://www.ibm.com/topics/machine-learning> (дата звернення 17.01.2024).
14. What is deep learning and how does it work? URL: <https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network> (дата звернення 12.12.2024).
15. What Is Deep Learning? How It Works, Techniques. URL: <https://www.mathworks.com/discovery/deep-learning.html> (дата звернення 12.12.2024).
16. ChatGPT від OpenAI на що він здатен та як ним скористатися? URL: https://www.imena.ua/blog/chatgpt/?gclid=EAIaIQobChMIp-mGo9f8gQMvVjIDBx22rAZYEAAYASAAEGJcnPD_BwE (дата звернення 15.12.2024).