

УДК: 512.7

*Родюк А. І., здобувачка вищої освіти,  
Луценко А. В., д-р філос. з математики,  
в. о. зав. кафедри прикладної математики та кібербезпеки,  
Донецький національний університет імені Василя Стуса*

## **ПРО МАТЕМАТИЧНІ АСПЕКТИ НЕЙРОННИХ МЕРЕЖ**

*Анотація. У роботі ідеться про математичні концепції нейронних мереж, про їх значення в сучасному світі. Проаналізовано значення лінійної алгебри, теорії ймовірності, теорії інформації та чисельних методів для глибокого навчання. На основі цього підкреслено важливість математики для підвищення якості життя, безпеки, стійкості та успішного майбутнього суспільства.*

*Ключові слова: нейронні мережі, лінійна алгебра, теорія ймовірностей, теорія інформації, чисельні методи, диференціальне числення, градієнтний спуск.*

**Вступ.** Нейронні мережі надзвичайно актуальні в сучасному світі через їх здатність вирішувати складні задачі в різних галузях. Вони використовуються для розпізнавання та класифікації об'єктів на зображеннях, що знаходить своє застосування у медицині та охоронних системах; обробки природної мови для перекладів, аналізу тексту, чат-ботів та голосових асистентів; прогнозування ринкових трендів, оцінки ризиків та виявлення шахрайства для фінансових послуг та ін. Нейронні мережі перетворюються в універсальний та широко використовуваний інструмент для різних сфер [1].

У традиційному підході до програмування ми говоримо комп'ютеру, що робити, розбиваючи великі проблеми на безліч маленьких точно визначених завдань, які він може легко виконати. В нейронній мережі ми не говоримо комп'ютеру, як вирішити нашу проблему. Замість цього він вчиться на даних, спостереженнях, знаходячи власне рішення поставленої проблеми. У 2006 році відбулося відкриття методів навчання в глибоких нейронних мережах, яке отримало назву глибоке навчання (deep learning) [2].

Метою роботи є демонстрація математики як фундаменту для функціонування нейронних мереж, а саме для глибокого навчання, що підкреслює її важливість у сучасному світі; ознайомлення з основними математичними концепціями, що лежать в основі нейронних мереж: лінійною алгеброю, теорією ймовірностей, статистикою, теорією інформації, чисельними методами та диференціальним численням, необхідними для навчання, оцінки й вдосконалення моделей нейронних мереж, адже в основі будь-якої нейронної мережі лежить розв'язання складних математичних задач.

**Основний текст. Лінійна алгебра.** Хороше розуміння основних понять лінійної алгебри, як-от скаляри, вектори, матриці й тензори, їх властивості та дії над ними, має важливе значення для досягнення та роботи з багатьма алгоритмами машинного навчання. Прикладом слугує метод головних компонент (principal components analysis PCA). Він застосовується для зменшення розмірності даних, витративши водночас найменшу кількість інформації. Він може бути зведений

до обчислення сингулярного розкладу матриці даних або обчислення власних векторів та чисел коваріаційної матриці початкових даних [1, 3].

**Теорія ймовірностей.** Майже всі види діяльності вимагають певної здатності міркувати в умовах невизначеності. Крім математичних тверджень, які є істинними за визначенням, важко уявити будь-яке абсолютно істинне судження, або будь-яку абсолютно ймовірну подію. Машинне навчання завжди має справу з невизначеними величинами, а невизначеність і стохастичність вимагають застосування теорії ймовірностей.

У програмах штучного інтелекту вона використовується двома основними способами. Перший: закони ймовірностей визначають міркування систем штучного інтелекту. Ймовірність можна розглядати як розширення логіки для роботи з невизначеністю. Логіка надає набір формальних правил для визначення того, які судження мають на увазі як істинні або хибні, якщо припустити, що деякий інший набір суджень є істинним або хибним. Теорія ймовірностей надає набір формальних правил для визначення ймовірності істинності судження за умови ймовірності інших суджень. Другий: теорію ймовірностей і статистику можна використовувати для теоретичного аналізу поведінки систем штучного інтелекту.

**Теорія інформації.** Теорія інформації також відіграє важливу роль у машинному навчанні, оскільки вона допомагає оцінювати кількість збереженої інформації для уникнення перенавчання; використовується для оптимізації способів кодування даних задля зменшення зберігання й передачі інформації; застосовується у визначенні ймовірностей, що є важливим для моделювання та прогнозування [1, 3].

**Чисельні методи.** Алгоритми машинного навчання зазвичай вимагають великої кількості чисельних методів. Особливо ті алгоритми, що вирішують математичні задачі за допомогою методів, які ітеративно оновлюють оцінки розв'язку, а не аналітично виводять формулу, що забезпечує символічний вираз для правильного рішення. До поширених операцій відносять розв'язування систем лінійних рівнянь і знаходження значення аргументу, що мінімізує функцію. Навіть проста оцінка математичної функції на цифровому комп'ютері може бути складною, коли функція включає дійсні числа, які не можуть бути точно представлені за допомогою кінцевого обсягу пам'яті. Однією з форм помилки є округлення чисел поблизу нуля до нуля та округлення великих чисел до  $\infty$ . Цих помилок можна уникнути, якщо виконати елементарні перетворення виразу [1, 3].

**Диференціальне числення.** Більшість алгоритмів глибокого навчання передбачають певну оптимізацію параметрів нейронних мереж. Це дає змогу моделі навчатися і вдосконалюватися, зменшуючи помилки у власних прогнозах. Під оптимізацією розуміється задача мінімізації або максимізації деякої функції  $f(x)$  шляхом збільшення її аргументу  $x$ . Для цих цілей використовується похідна  $f'(x)$ , яка показує, як потрібно змінити аргумент  $x$  для того, щоб отримати задане покращення функції  $f(x)$ . Наприклад, ми знаємо, що  $f(x - \varepsilon f'(x))$  менше ніж  $f(x)$  для досить малих  $\varepsilon$ . Отже, можна зменшити  $f(x)$ , переміщуючи  $x$  маленькими кроками з протилежним знаком похідної. Ця техніка називається градієнтним спуском.

**Висновки.** Нейронні мережі відіграють ключову роль у сучасних технологіях і мають величезне значення для багатьох галузей. Їх розвиток сприяє інноваціям та прогресу технологій майбутнього, як-от штучний інтелект і розумні системи; покращенню якості життя завдяки створенню нових медичних інструментів для діагностики; покращенню охоронних систем та персоналізованих рекомендацій; автоматизації складних завдань, зменшуючи потребу в ручній праці та покращуючи результати; аналізу і розумінню великих обсягів даних; точнішим прогнозуванням ринкових трендів, природних явищ та інших важливих подій.

Гарне розуміння основних математичних концепцій, що лежать у їх основі, як-от лінійна алгебра, теорія ймовірності, теорія інформації, чисельні методи та диференціальне числення дають змогу краще зрозуміти й покращити ефективність і точність механізмів функціонування, процесів оптимізації та навчання нейронних мереж.

Покладаючись на неоціненне значення нейронних мереж для суспільства та їх побудові на математиці, можна зазначити, що математика є необхідним фундаментальним інструментом для побудови успішного майбутнього людства.

#### Список використаних джерел

1. Nielsen M. A. *Neural Networks and Deep Learning*. San Francisco: Determination Press, 2015. 431 p.
2. Bishop C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006. 738 p.
3. Goodfellow I., Bengio Y., Courvill A. *Deep Learning*. Cambridge, Massachusetts: MIT Press, 2016. 800 p.