

ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ЯК МАТЕМАТИЧНИХ МОДЕЛЕЙ

Військовий інститут телекомунікацій та інформатизації
імені Героїв Крут

Анотація

У роботі розглянуто нейронні мережі як сучасний інструмент математичного моделювання складних процесів і явищ. Досліджено математичні основи побудови штучних нейронних мереж, принципи їх функціонування та навчання. Проаналізовано роль методів лінійної алгебри, математичного аналізу та оптимізації у формуванні моделей на основі нейронних мереж. Висвітлено можливості застосування нейронних мереж для апроксимації функцій, аналізу даних, прогнозування та розв'язування прикладних задач. Показано, що нейронні мережі є ефективним засобом моделювання складних нелінійних залежностей і знаходять широке застосування в сучасних інформаційних технологіях.

Ключові слова: нейронна мережа, математичне моделювання, штучний інтелект, оптимізація, апроксимація функцій, машинне навчання, математична модель.

Abstract

This paper considers neural networks as a modern tool for mathematical modeling of complex processes and phenomena. The mathematical foundations of artificial neural network construction, as well as the principles of their operation and training, are investigated. The role of linear algebra, mathematical analysis, and optimization methods in the development of neural network-based models is analyzed. The possibilities of applying neural networks to function approximation, data analysis, forecasting, and solving applied problems are highlighted. It is shown that neural networks are an effective means of modeling complex nonlinear relationships and are widely used in modern information technologies.

Keywords: neural network, mathematical modeling, artificial intelligence, optimization, function approximation, machine learning, mathematical model.

Вступ

Сучасний розвиток інформаційних технологій супроводжується широким використанням методів штучного інтелекту, зокрема нейронних мереж. Вони застосовуються для аналізу даних, прогнозування, розпізнавання образів та прийняття рішень. З математичної точки зору нейронні мережі є складними нелінійними моделями, що дозволяють апроксимувати багатовимірні функції та описувати складні залежності між змінними.

Мета роботи полягає в аналізі нейронних мереж як математичних моделей, дослідженні їх структури, принципів навчання та основних сфер застосування.

Результати дослідження

Математична модель штучного нейрона

Нейронні мережі є одним із найперспективніших напрямів сучасного математичного моделювання та штучного інтелекту. Їх широке застосування зумовлене здатністю ефективно обробляти великі обсяги даних, виявляти приховані закономірності та прогнозувати результати складних процесів [1, 2]. Сьогодні нейронні мережі використовуються у фінансовій сфері, медицині, робототехніці, системах підтримки прийняття рішень та багатьох інших галузях.

Математичною основою нейронної мережі є штучний нейрон, який моделює принцип роботи біологічного нейрона. Його функціонування можна описати формулою:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

де x_i – вхідні дані, w_i – вагові коефіцієнти, b – зміщення, f – функція активації.

Функція активації визначає, як нейрон перетворює отриманий сигнал та передає його далі. Саме вона забезпечує нелінійність моделі, що дозволяє нейронним мережам розв'язувати складні задачі.

Функція активації вводить нелінійність у модель, що дозволяє мережі описувати складні залежності. Найпоширенішими є **sigmoid**, **ReLU** та **tanh**.

- sigmoid (сигмоїдна функція)- перетворює будь-яке вхідне значення в число від 0 до 1.
- ReLU (Rectified Linear Unit)- найпоширеніша функція активації в сучасних нейронних мережах.
- tanh (гіперболічний тангенс) - функція схожа на sigmoid, але її значення лежать у діапазоні від -1 до 1 .

У сучасних моделях штучного інтелекту найчастіше використовується **ReLU**, оскільки вона забезпечує швидше та ефективніше навчання нейронних мереж. Для вихідних шарів класифікаційних моделей часто застосовують **sigmoid**, а **tanh** використовується в окремих архітектурах рекурентних нейронних мереж.

Таким чином, нейрон виконує операцію зваженого підсумовування з подальшим нелінійним перетворенням.

Нейронна мережа як композиція функцій

Нейронна мережа являє собою систему з багатьох шарів нейронів, де кожен шар виконує перетворення даних:

$$f(x) = f^3 f^2 f^1(x),$$

Така структура дозволяє моделювати складні нелінійні функції. З математичної точки зору нейронна мережа є композицією багатьох векторно-матричних перетворень.

Навчання нейронної мережі як задача оптимізації

Нейронна мережа складається з множини взаємопов'язаних нейронів, організованих у вхідний, приховані та вихідний шари. З математичної точки зору вона є композицією багатьох функцій, що дозволяє описувати складні нелінійні залежності між вхідними та вихідними даними [3, 4]. Саме ця властивість забезпечує високу ефективність нейронних мереж під час розв'язування задач класифікації, прогнозування та апроксимації функцій.

Важливим етапом функціонування нейронної мережі є процес навчання, який полягає у підборі оптимальних значень вагових коефіцієнтів. Математично ця задача зводиться до мінімізації функції втрат за допомогою методів оптимізації [4]. Одним із найпоширеніших алгоритмів навчання є градієнтний спуск:

Процес навчання полягає у підборі таких значень ваг w_i , які мінімізують функцію втрат:

$$L(w) \rightarrow \min,$$

Для цього використовується метод градієнтного спуску:

$$w := w - \eta \cdot \nabla L(w),$$

де η - швидкість навчання, $\nabla L(w)$ – градієнт функції втрат.

Таким чином, навчання нейронної мережі зводиться до багатовимірної задачі оптимізації.

Нейронні мережі як апроксиматори функцій

Однією з найважливіших властивостей нейронних мереж є здатність до апроксимації функцій. Згідно з теоремою універсальної апроксимації, багат шарова нейронна мережа може наблизити довільну неперервну функцію із заданою точністю [3]. Це робить нейронні мережі ефективним інструментом математичного моделювання процесів, для яких складно або неможливо побудувати точну аналітичну модель.

Математичне представлення даних

У нейронних мережах дані подаються у вигляді векторів, а обчислення — у вигляді матричних операцій:

$$Y = WX + b$$

де W – матриця ваг, X – вектор вхідних даних.

Це дозволяє ефективно реалізовувати обчислення за допомогою лінійної алгебри.

Сучасні нейронні мережі активно використовуються для розпізнавання образів, аналізу текстової інформації, прогнозування часових рядів та обробки великих масивів даних [2, 6]. Їх застосування дозволяє отримувати результати високої точності навіть у випадках, коли досліджувані процеси мають складну структуру та значну кількість параметрів.

Висновки

Нейронні мережі є потужним математичним інструментом моделювання складних нелінійних систем. Їхня основа базується на методах лінійної алгебри, математичного аналізу та теорії оптимізації. Завдяки здатності до апроксимації функцій та навчанню на даних вони знаходять широке застосування в сучасних інформаційних технологіях.

Таким чином, нейронні мережі є потужним математичним інструментом, який поєднує методи лінійної алгебри, математичного аналізу, теорії ймовірностей та оптимізації. Завдяки своїй універсальності та високій ефективності вони посідають важливе місце серед сучасних засобів математичного моделювання і залишаються одним із найбільш перспективних напрямів розвитку інформаційних технологій.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Artificial Intelligence: A Modern Approach / Russell S., Norvig P. 4th ed. Pearson, 2021. 1136 p.
2. Neural Networks and Deep Learning. Determination Press, 2015. 225 p.
3. Deep Learning / Goodfellow I., Bengio Y., Courville A. MIT Press, 2016. 775 p.
4. Pattern Recognition and Machine Learning. Springer, 2006. 738 p.
5. Machine Learning. McGraw-Hill, 1997. 414 p.
6. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 3rd ed. O'Reilly Media, 2022. 851 p.

Абламська Валентина Михайлівна викладач, Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, м. Київ, email: boltyk@ukr.net

Ablamska Valentyna lecturer, Kruty Heroes Military Institute of Telecommunications and Information Technology, Kyiv, email: boltyk@ukr.net