

# АЛГОРИТМИ В МАШИННОМУ НАВЧАННІ

<sup>1</sup>Вінницький національний технічний університет

## Анотація

У роботі розглянуто основні алгоритми машинного навчання та їх роль у сучасних інформаційних технологіях. Проведено класифікацію алгоритмів за типами навчання: з учителем, без учителя та з підкріпленням. Проаналізовано принципи роботи найпоширеніших алгоритмів, таких як лінійна регресія, дерева рішень та метод *k*-середніх. Визначено їх переваги, недоліки та сфери застосування. Показано, що ефективність алгоритмів значною мірою залежить від якості даних та вибору моделі.

**Ключові слова:** машинне навчання, алгоритми, класифікація, регресія, кластеризація, аналіз даних.

## Abstract

The paper considers the main algorithms of machine learning and their role in modern information technologies. A classification of algorithms by learning types is presented: supervised, unsupervised, and reinforcement learning. The principles of operation of the most common algorithms, such as linear regression, decision trees, and *k*-means, are analyzed. Their advantages, disadvantages, and areas of application are identified. It is shown that the efficiency of algorithms largely depends on data quality and model selection.

**Keywords:** machine learning, algorithms, classification, regression, clustering, data analysis.

## Вступ

Машинне навчання є одним із найважливіших напрямів розвитку сучасної інформатики, який дозволяє створювати системи, здатні навчатися на основі даних [1]. Основна ідея полягає у побудові функції

$$y=f(x,\theta), \quad (1)$$

де  $x$  — вхідні дані,  $y$  — прогнозований результат,  $\theta$  — параметри моделі.

Метою навчання є підбір таких параметрів  $\theta$ , які мінімізують функцію втрат:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

## Результати дослідження

Алгоритми машинного навчання поділяються на три основні категорії: навчання з учителем, без учителя та з підкріпленням.

Навчання з учителем передбачає використання розмічених даних, де кожному прикладу відповідає правильна відповідь [2]. Основною задачею є побудова функції вигляду:

$$y=f(x,\theta), \quad (3)$$

де  $x$  — вхідні дані,  $y$  — вихід,  $\theta$  — параметри моделі.

Одним із базових алгоритмів є лінійна регресія, яка використовується для прогнозування числових значень:

$$\hat{y}=w_0+w_1x. \quad (4)$$

Навчання моделі здійснюється шляхом мінімізації функції втрат:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

що дозволяє знайти оптимальні параметри. Для цього часто використовується метод градієнтного спуску:

$$\theta := \theta - \alpha \nabla L(\theta), \quad (6)$$

де  $\alpha$  — швидкість навчання.

Для задач класифікації застосовуються алгоритми, такі як дерева рішень та логістична регресія. Ймовірність належності об'єкта до класу визначається за допомогою логістичної функції:

$$P(y=1|x) = \frac{1}{1+e^{-z}}, \quad z = w^T x. \quad (7)$$

Навчання без учителя працює з нерозміченими даними та спрямоване на виявлення прихованих структур. Одним із найпоширеніших алгоритмів є метод k-середніх, який використовується для кластеризації. Його метою є мінімізація функції:

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (8)$$

де  $\mu_i$  — центр кластера, а  $C_i$  — множина об'єктів у кластері.

Навчання з підкріпленням базується на взаємодії алгоритму із середовищем. Модель отримує винагороду або штраф залежно від виконаних дій і намагається максимізувати сумарну винагороду [3]. Одним із ключових підходів є Q-навчання:

$$Q(s,a) = r + \gamma \max_{a'} Q(s',a'), \quad (9)$$

де  $r$  — винагорода,  $\gamma$  — коефіцієнт дисконту.

Важливим аспектом є оцінка якості моделей. Для цього використовуються метрики, такі як точність (accuracy), повнота (recall) та F-міра. Крім того, необхідно враховувати проблему перенавчання, яка виникає тоді, коли модель надто точно підлаштовується під навчальні дані і втрачає здатність до узагальнення [4].

Таким чином, різні типи алгоритмів машинного навчання мають свої особливості та сфери застосування. Вибір конкретного методу залежить від типу задачі, структури даних і вимог до точності результатів.

## Висновки

У результаті дослідження встановлено, що алгоритми машинного навчання є ефективним інструментом для аналізу даних і прогнозування. Використання математичних моделей та оптимізаційних методів дозволяє підвищити точність і надійність результатів. Подальший розвиток алгоритмів машинного навчання сприятиме створенню інтелектуальних систем у різних галузях.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Mitchell T. Machine Learning. – McGraw-Hill, 1997.
2. Bishop C. Pattern Recognition and Machine Learning. – Springer, 2006.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press, 2016..
4. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. – Springer, 2009.

**Балан Владислав Русланович** — студент групи ІКІ-25МС, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail:

[vladislavbalan916@gmail.com](mailto:vladislavbalan916@gmail.com)

Науковий керівник: **Добровольська Наталія Вікторівна** – к.пед.н., доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, Вінниця, e-mail: [do\\_b\\_r\\_n\\_v@vntu.edu.ua](mailto:do_b_r_n_v@vntu.edu.ua)

**Balan Vladislav Ruslanovich** — student of group 1KI-25MS, Faculty of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: [vladislavbalan916@gmail.com](mailto:vladislavbalan916@gmail.com)

Supervisor: **Dobrovolska Nataliia V.** – Ph.D., Associate Professor of the Computer Engineering Department, Vinnytsia National Technical University, Vinnytsia, e-mail: [do\\_b\\_r\\_n\\_v@vntu.edu.ua](mailto:do_b_r_n_v@vntu.edu.ua)