

## **МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ ДІАГНОСТУВАННЯ СТАНУ ЗДОРОВ'Я ЛЮДИНИ**

Вінницький національний технічний університет

### ***Анотація***

*У сучасному світі інформаційні технології набувають все більшого значення у сфері охорони здоров'я, оскільки вони здатні вирішувати низку критично важливих завдань, пов'язаних із забезпеченням точності та швидкості діагностування. Вибір найбільш точного алгоритму є критично важливим для досягнення успішних результатів. У роботі було досліджено та проаналізовано різні алгоритми такі як: градієнтний бустинг дерев рішень, випадковий ліс, дерева рішень з метою розробки системи діагностування стану здоров'я людини.*

**Ключові слова:** системи діагностування, здоров'я, точність.

### ***Abstract***

In today's world, information technologies are becoming more and more important in the field of health care, as they are able to solve a number of critically important tasks related to ensuring the accuracy and speed of diagnosis. Choosing the most accurate algorithm is critical to achieving successful results. In the work, various algorithms were researched and analyzed, such as: gradient boosting of decision trees, random forest, decision trees for the purpose of developing a system for diagnosing the state of human health.

**Keywords:** diagnostic systems, health, accuracy.

### **Вступ**

У сучасному світі діагностування здоров'я стає все більш важливим і отримує все більше конкурентних рішень. З розвитком технологій штучного інтелекту проблема діагностування здоров'я отримала нові рішення, які значно покращують точність діагнозу та спрощення самого процесу діагностування. Одним з рішень є діагностування стану здоров'я за допомогою алгоритмів машинного навчання [1].

Ця робота спрямована на аналіз різних методів для створення систем діагностування здоров'я людини, зокрема з використанням алгоритмів машинного навчання. Завдяки застосуванню штучного інтелекту та машинного навчання, з'являється можливість підвищення рівня обробки та аналізу великих обсягів медичних даних, що дозволяє значно покращити діагностичні процедури, а також виявляти захворювання на ранніх стадіях. Необхідність швидкого і точного діагностування стану пацієнта в області охорони здоров'я потребує залучення новітніх технологій для обробки інформації та створення систем прогнозування.

Метою роботи є дослідження та аналіз існуючих методів машинного навчання для діагностування стану здоров'я людини, що забезпечить більш швидкі та точні результати для пацієнтів.

### **Результати дослідження**

Для створення системи діагностування можуть бути використані різні методи машинного навчання. Метод градієнтного бустингу дерев рішень є одним із найпопулярніших й ефективних підходів у машинному навчанні для вирішення завдань класифікації та регресії. Його суть полягає у побудові ансамблю слабких моделей, якими зазвичай виступають дерева рішень, які працюють послідовно, щоб покращити точність прогнозування. Даний метод широко застосовується для роботи з великими масивами даних, особливо у задачах, де необхідна висока точність прогнозування. Зокрема, у медицині цей метод можна використовувати для прогнозування стану здоров'я пацієнтів, оцінки ризику розвитку різних захворювань та діагностики на основі історичних даних [2].

Метод випадкового лісу є одним із методів ансамблевого навчання, який базується на використанні множини дерев рішень. Ідея полягає в тому, щоб побудувати велику кількість дерев рішень на основі різних випадкових підмножин даних і ознак, а потім об'єднати результати кожного дерева, використовуючи голосування для задач класифікації або усереднення для задач регресії. У медичних додатках випадковий ліс може бути корисним для класифікації пацієнтів за групами ризику, прогнозування результатів лікування та оцінки ймовірності розвитку певних захворювань. Завдяки своїй здатності ефективно працювати з великою кількістю ознак і обробляти пропущені значення, випадковий ліс може допомогти виявляти складні закономірності в медичних даних [3].

Метод дерева рішень є одним із найпоширеніших підходів у машинному навчанні, який використовується для розв'язання задач класифікації та регресії. Дерева рішень створюють модель у вигляді структури, що нагадує дерево, де кожен вузол відповідає за перевірку певної ознаки, а кожна гілка – за результат цієї перевірки. У медичній сфері дерева рішень можуть використовуватися для прогнозування розвитку захворювань або визначення груп ризику [4].

Для вдосконалення роботи алгоритмів необхідно врахувати наступні фактори: наявність захворювань серця, артеріальний тиск, рівень холестерину, загальний стан здоров'я, вікова група, діабет, куріння, індекс маси тіла, рівень фізичної активності, харчування. Інтеграція цих факторів допоможе більш точно діагностувати стан здоров'я людини відповідно до її особистих атрибутів. Також необхідно забезпечити конфіденційність персональної інформації користувачів на кожному етапі розробки та впровадження системи діагностування стану здоров'я людини.

### Висновки

Актуальним є дослідження алгоритмів машинного навчання для діагностування стану здоров'я людини, що дозволяє намітити шляхи розвитку і вдосконалення методів штучного інтелекту. У результаті дослідження було розглянуто такі методи машинного навчання як метод градієнтного бустингу дерев, випадковий ліс та дерева рішень. Вибір методів здійснюється залежно від потреб проекту. Впровадження методів машинного навчання для діагностування стану здоров'я людини дозволяє значно підвищити точність та швидкість результатів.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Штучний інтелект. Машинне навчання / Григоров О. В., Аніщенко Г. О., Стрижак В. В., Петренко Н. О., Турчин О. В., Окунь А. О., Пономарьов О. Е. Сучасні технології : зб. наук. пр. / М-во освіти і науки України, ХНАДУ. - Харків, 2019. - 15 с.
2. Alcolea, A., & Resano, J. FPGA accelerator for gradient boosting decision trees. *Electronics*. 2021. Vol 10, No 3, 314 p.
3. Schonlau, M., & Zou, R. Y. The random forest algorithm for statistical learning. *The Stata Journal*. 2020. Vol 20, No 1, pp. 3-30. *Machine Learning Basics: Random Forest Classification*. URL: <https://towardsdatascience.com/machine-learning-basics-random-forest-classification-499279bac51e> (дата звернення 06.11.2024).
4. Why Do We Use Decision Trees in Machine Learning? URL: <https://www.turing.com/kb/importance-of-decision-trees-in-machine-learning> (Дата звернення 06.11.2024).

**Пасічний Іван Вікторович** – студент групи 2КН-23м, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: pasichnyuk1994@gmail.com.

**Хоменко Володимир Михайлович** – студент групи ICT-23б, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: vova.khomenko2100@gmail.com.

**Іванчук Ярослав Володимирович** – д.т.н., професор комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця, e-mail: ivanchuck@ukr.net.

**Pasichniuk Ivan V.** – student of the 2KN-23m group, faculty of intellectual information technologies and automation, Vinnytsia National Technical University, Vinnytsia, e-mail: pasichnyuk1994@gmail.com.

**Khomenko Volodymyr M.** – student of the 1IST-23b group, faculty of intellectual information technologies and automation, Vinnytsia National Technical University, Vinnytsia, e-mail: vova.khomenko2100@gmail.com.

**Ivanchuk Yaroslav V.** – Dr. Sc. (Eng.), Professor, Department of Computer Science, Vinnytsia National Technical University, Vinnytsia, e-mail: ivanchuck@ukr.net.