

ДІАГНОСТИКА ЗАХВОРЮВАНЬ ПО АНАЛІЗАМ КРОВІ ЗАСОБАМИ ШІ

Зозуля Н.О., Гриньова О.Є.

e-mail: nikita.zozulia1@nure.ua, olena.hrynova@nure.ua

Харківський національний університет радіоелектроніки, каф. ШІ
м. Харків, Україна

With the expansion of artificial intelligence integration in medicine, the development of systems capable of automatically diagnosing diseases based on blood test data becomes particularly relevant. The main challenge lies in collecting and standardizing diverse data (language, measurement units, reference norms for different age groups). The goal of this research is to create a system capable of aggregating, normalizing, and transmitting data for AI training.

Останні досягнення в галузі штучного інтелекту суттєво змінили медицину, покращуючи діагностику, прогнозування та персоналізоване лікування. Такі технології, як глибинне навчання та алгоритми машинного навчання, вже демонструють високу ефективність у виявленні онкологічних, кардіологічних та неврологічних захворювань. Автоматизовані системи аналізу медичних зображень, електронних записів пацієнтів та лабораторних тестів значно скорочують час постановки діагнозу та мінімізують ризик лікарських помилок.

Одним із найперспективніших напрямків є аналіз біохімічних та загальноклінічних показників крові за допомогою ШІ [1]. Велика кількість параметрів, їхня варіативність залежно від віку, статі, регіону проживання та медичних стандартів створюють суттєві виклики при розробці такої системи. Крім того, методи забору та аналізу крові можуть відрізнятися, що вимагає уніфікації отриманих даних. Для побудови моделей, здатних розрізняти різні захворювання на основі показників крові, застосовуються такі алгоритми машинного навчання, як Support Vector Machines (SVM), Random Forest та логістична регресія.

Основні виклики під час підготовки даних для навчальної вибірки:

– стандартизація мовних даних. Медичні показники можуть надаватися різними мовами, що ускладнює їхню автоматичну обробку. Оптимальним варіантом на даному етапі буде використання обробки природної мови (NLP) – механізмів для автоматичного перекладу та кодування медичних показників, представлених на різних мовах та вилучення важливої інформації про пацієнтів [2];

– уніфікація одиниць вимірювання. Різні лабораторії використовують відмінні одиниці вимірювання для тих самих параметрів (наприклад, г/л проти мг/дл). Це вимагає впровадження єдиної системи перерахунку, що включає: автоматичне розпізнавання одиниць вимірювання, конвертація

одиниць у стандартний формат та уніфікація через медичні бази знань (наприклад, LOINC) [3];

– нормативні значення. У різних вікових груп, представників різних статей і навіть у різних країнах можуть бути свої нормативи для одних і тих самих показників. Система має враховувати ці відмінності;

– автоматизований збір та обробка даних. Запропонована система повинна мати механізми для автоматичного збору інформації з електронних медичних записів (EMR) та лабораторних інформаційних систем (LIS), а також її подальшої обробки для забезпечення точності [4].

Окрім збору, стандартизації та уніфікації медичних даних, ключовим завданням є створення моделі, здатної не лише обробляти показники аналізів, але й встановлювати діагнози на їх основі. Для цього використовується глибинне навчання (Deep Learning), що дозволяє знаходити складні закономірності між лабораторними показниками, симптомами пацієнта та кінцевими діагнозами. Найефективнішими методами в цьому контексті є згорткові нейронні мережі (CNN) для обробки зображень (наприклад, сканованих аналізів), рекурентні нейронні мережі (RNN) та трансформери, а також глибокі багатошарові перцептрони (MLP), що можуть навчатися на структурованих числових медичних даних. Навчання таких моделей відбувається на великому масиві попередньо підготовлених медичних даних, що дозволяє їм прогнозувати ймовірні діагнози на основі отриманих лабораторних тестів та скарг пацієнта, підвищуючи точність та швидкість діагностики.

Використання глибинного навчання дозволить значно підвищити точність діагностики, зменшити час на інтерпретацію результатів аналізів та сприяти ранньому виявленню захворювань. Медичні лабораторії зможуть автоматизувати та оптимізувати процес аналізу даних. Інтелектуальний аналіз великих обсягів стандартизованих даних сприятиме подальшому розвитку медичної науки у виявленні залежностей параметрів аналізу крові.

Запропонований підхід дозволить створити комплексну базу даних для навчання інтелектуальних діагностичних систем, що сприятиме підвищенню точності діагностики захворювань на основі аналізу крові.

Список використаних джерел:

1. Дудник М.П., Удовенко С.Г., Чала Л.Е., Соколовська М.М. Нейромережева технологія багатомовної класифікації електронних текстів // Біоніка інтелекту. – 2021. – Вип. 2 (97). – С. 3-12
2. Topol, E. (2019). Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. Basic Books.
3. WHO (Всесвітня організація охорони здоров'я). Laboratory Quality Standards and Guidelines, 2022.
4. Shilo S., Rossman H., Segal E. Axes of a revolution: challenges and promises of big data in healthcare. Nature Medicine, 2020.