

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Центр \_\_\_\_\_ ННЦЗФН \_\_\_\_\_  
(повна назва)

Кафедра \_\_\_\_\_ Штучного інтелекту \_\_\_\_\_  
(повна назва)

## КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти \_\_\_\_\_ перший (бакалаврський) \_\_\_\_\_

\_\_\_\_\_ AI-орієнтована діагностична система «HemaIntel» \_\_\_\_\_  
(тема)

Виконав:  
здобувач \_\_\_\_\_ четвертого \_\_\_\_\_ року навчання,  
групи \_\_\_\_\_ ІТШЗ-21-1 \_\_\_\_\_

\_\_\_\_\_ Нікіта Зозуля \_\_\_\_\_  
(власне ім'я, прізвище)

Спеціальність \_\_\_\_\_ 122 Комп'ютерні науки \_\_\_\_\_  
(код і повна назва спеціальності)

Тип програми \_\_\_\_\_ освітньо-професійна \_\_\_\_\_  
Освітня програма \_\_\_\_\_ Штучний інтелект \_\_\_\_\_  
(повна назва освітньої програми)

Керівник \_\_\_\_\_ ст. викл. Олена Гриньова \_\_\_\_\_  
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ \_\_\_\_\_  
(підпис)

\_\_\_\_\_ Олег ЗОЛОТУХІН \_\_\_\_\_  
(власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Центр \_\_\_\_\_ ННЦЗФН \_\_\_\_\_

Кафедра \_\_\_\_\_ Штучного інтелекту \_\_\_\_\_

Рівень вищої освіти \_\_\_\_\_ перший (бакалаврський) \_\_\_\_\_

Спеціальність \_\_\_\_\_ 122 Комп'ютерні науки \_\_\_\_\_  
(код і повна назва)

Тип програми \_\_\_\_\_ освітньо-професійна \_\_\_\_\_

Освітня програма \_\_\_\_\_ Штучний інтелект \_\_\_\_\_  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_

(підпис)

«\_\_\_\_\_» \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ**  
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві \_\_\_\_\_ Зозулі Нікіті Олександровичу \_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема роботи \_\_\_\_\_ AI-орієнтована діагностична система «HemaIntel» \_\_\_\_\_

затверджена наказом університету від 07 травня 2025 р. № 80Стз

2. Термін подання студентом роботи до екзаменаційної комісії 24 червня 2025 р.

3. Вихідні дані до роботи \_\_\_\_\_ Науково-технічні публікації, дані Інтернет-джерел та відомих наукових проєктів, документації моделей від OpenAI, документації з мови програмування Python. \_\_\_\_\_

4. Перелік питань, що потрібно опрацювати в роботі \_\_\_\_\_

1) Аналіз предметної галузі


2) Опис та порівняльний аналіз існуючих методів та технологій


3) Опис програмної реалізації

## КАЛЕНДАРНИЙ ПЛАН

| №  | Назва етапів роботи                         | Строк / терміни виконання етапів роботи | Примітка |
|----|---|---|----------|
| 1  | Отримання завдання на кваліфікаційну роботу | 05.05.2025                              | виконано |
| 2  | Аналіз предметної галузі                    | 07.05.2025                              | виконано |
| 3  | Огляд існуючих методів та технологій        | 10.05.2025                              | виконано |
| 4  | Проектування системи                        | 13.05.2025                              | виконано |
| 5  | Програмна реалізація застосунку             | 02.06.2025                              | виконано |
| 6  | Написання пояснювальної записки             | 13.06.2025                              | виконано |
| 7  | Перевірка на академічний плагіат            | 16.06.2025                              | виконано |
| 8  | Нормоконтроль                               | 17.06.2025                              | виконано |
| 9  | Підготовка презентації та доповіді          | 18.06.2025                              | виконано |
| 10 | Попередній захист                           | 19.06.2025                              | виконано |
| 11 | Рецензування                                | 20.06.2025                              | виконано |
| 12 | Захист перед ЕК                             | 24.06.2025                              | виконано |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |
|    |   |   |          |

Дата видачі завдання 5 травня 2025 р.

Здобувач   
(підпис)

Керівник роботи   
(підпис)

ст. викл. Олена Гриньова  
(посада, власне ім'я, прізвище)

## РЕФЕРАТ

Пояснювальна записка: 66 с., 10 рис., 1 дод., 21 джерело.

АНАЛІЗ ДАНИХ, ДІАГНОСТИКА ЗАХВОРЮВАНЬ, МАШИННЕ НАВЧАННЯ, МЕДИЧНІ ЗОБРАЖЕННЯ, РЕНТГЕН, ШТУЧНИЙ ІНТЕЛЕКТ.

Тема роботи є актуальною через зростання потреби в ефективних інструментах для ранньої діагностики захворювань і підвищення доступності медичної допомоги за рахунок технологій штучного інтелекту.

Об'єкт дослідження – системи штучного інтелекту для діагностики захворювань за медичними даними.

Предмет дослідження – методи і моделі машинного навчання для прогнозування діагнозів на основі медичних даних.

Мета роботи – розробка інтелектуальної системи для діагностики захворювань на основі медичних аналізів, що дозволяє прогнозувати наявність патологій з використанням датасетів і забезпечує зручний інтерфейс для взаємодії користувача з системою. В результаті роботи очікується створення функціонального та інтуїтивно зрозумілого додатку, який дозволить отримувати рекомендації щодо стану здоров'я шляхом аналізу медичних даних.

Методи дослідження – аналіз наукової літератури та публічних досліджень у галузі медичної діагностики, методи машинного навчання, моделювання на основі відкритих даних.

## **ABSTRACT**

Bachelor's thesis contains: 66 pp., 10 fig., 1 ann., 21 references.

**DATA ANALYSIS, DISEASE DIAGNOSTICS, MACHINE TRAINING,  
MEDICAL IMAGING, PIECE INTELLIGENCE, X-RAY.**

The topic of the work is relevant due to the increasing need for effective tools for early diagnostic research and increasing the availability of medical services for calculating technological piece intelligence.

Reserch Object – is the systems of piece intelligence for diagnostic research on medical data.

Reserch Subject – is the methods and models of machine learning for predicting diagnoses based on new medical data.

Purpose of the work is the development of intelligent systems for diagnostic research of basic medical tests, which allow predicting the occurrence of pathologies based on local datasets and closing manual interfaces for interaction with the system. As a result of work on the creation of a functional and intuitive add-on that will allow you to select recommendations that will become a healthy path for analyzing medical data.

Research methods – analysis of new literature and published studies in the field of medical diagnostics, machine learning methods, modeling based on new critical data.

## ЗМІСТ

|  |    |
|--|----|
| Перелік умовних позначень, символів, одиниць, скорочень і термінів .....                           | 8  |
| Вступ.....   | 9  |
| 1 Аналіз предметної галузі .....   | 11 |
| 1.1 Історія розвитку технологій медичної діагностики за допомогою штучного інтелекту .....         | 11 |
| 1.2 Типи медичної діагностики та види даних, які використовуються в інтелектуальних системах ..... | 12 |
| 1.2.1 Фізикальні методи діагностики .....  | 12 |
| 1.2.2 Лабораторна діагностика .....  | 13 |
| 1.2.3 Інструментальна діагностика.....   | 15 |
| 1.2.4 Функціональна діагностика .....  | 17 |
| 1.3 Актуальність задачі медичної діагностики методами штучного інтелекту.....                      | 18 |
| 1.4 Мета роботи та постановка задачі.....  | 20 |
| 1.4.1 Мета роботи.....   | 20 |
| 1.4.2 Постановка задачі .....  | 21 |
| 2 Огляд та порівняльний аналіз існуючих методів та технологій .....                                | 22 |
| 2.1 Сучасний стан розробок у галузі медичних AI-систем.....  | 22 |
| 2.2 Огляд існуючих програмних рішень.....  | 23 |
| 2.2.1 CheXNet.....   | 23 |
| 2.2.2 DeepMind Health .....  | 25 |
| 2.2.3 Watson for Oncology.....   | 27 |
| 2.3 Методи, що використовуються в існуючих програмних рішеннях... ..                               | 28 |
| 2.3.1 Глибоке навчання для медичних зображень.....   | 28 |
| 2.3.2 Машинне навчання .....   | 31 |
| 2.3.3 Часові ряди .....  | 34 |
| 2.3.4 Експертні системи.....   | 36 |
| 2.3.5 Генетичні алгоритми .....  | 38 |

|  |    |
|--|----|
| 2.3.6 Медичні системи на основі NLP .....                | 39 |
| 2.3.7 Медичні системи на основі LLM .....                | 41 |
| 2.3.8 Нейроінтерфейс «мозок-машина».....                 | 43 |
| 2.3.9 Імітаційне моделювання лікарських препаратів ..... | 45 |
| 3 Опис програмної реалізації .....                       | 48 |
| 3.1 Frontend .....                                       | 49 |
| 3.1.1 HTML .....   | 49 |
| 3.1.2 CSS.....   | 49 |
| 3.1.3 JavaScript.....                                    | 50 |
| 3.2 Backend.....   | 50 |
| 3.2.1 Арі.....   | 51 |
| 3.2.2 Навчання моделей.....                              | 51 |
| 3.3 Робота з датасетами .....                            | 53 |
| 3.4 База даних .....                                     | 54 |
| 3.5 Логіка роботи системи.....                           | 57 |
| 4 Інтерфейс користувача .....                            | 58 |
| Висновки .....   | 61 |
| Перелік джерел посилання .....                           | 63 |
| Додаток А Відомість кваліфікаційної роботи .....         | 66 |

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

АЛТ – аланінамінотрансфераза;

АСТ – аспартатамінотрансфераза;

ЕКГ – електрокардіограма;

КТ – комп'ютерна томографія;

МРТ – магнітно-резонансна томографія;

ТТГ – тиреотропний гормон;

УЗД – ультразвукове дослідження;

AI – Artificial Intelligence – штучний інтелект;

API – Application Programming Interface – програмний інтерфейс  
прикладного рівня;

CNN – Convolutional Neural Network – згортова нейронна мережа;

CSV – Comma-Separated Values – формат файлів для зберігання  
табличних даних;

DL – Deep Learning – глибинне навчання;

JSON – JavaScript Object Notation – формат обміну даними;

ML – Machine Learning – машинне навчання;

REST – Representational State Transfer – архітектурний стиль взаємодії  
компонентів у вебзастосунках;

SQLite – легка реляційна система керування базами даних.

## ВСТУП

Сучасна медицина переживає стрімкий розвиток, що супроводжується не лише появою нових методів діагностики та лікування, але й суттєвим збільшенням обсягу даних, які повинен опрацьовувати лікар під час прийняття клінічних рішень. Обсяг інформації про пацієнтів – від результатів лабораторних аналізів до медичних зображень та електрокардіограм – зростає в геометричній прогресії, що значно ускладнює роботу медичних працівників і підвищує ризики помилок через людський фактор. У цих умовах особливої актуальності набувають інтелектуальні системи, здатні автоматично аналізувати дані та допомагати у постановці діагнозу.

Однією з головних проблем сучасної охорони здоров'я є зростання навантаження на медичних працівників, особливо у країнах із обмеженими ресурсами та значним дефіцитом медичного персоналу. Це призводить до того, що лікарі часто змушені приймати клінічні рішення в умовах браку часу й необхідності одночасного аналізу великої кількості гетерогенних даних. Використання штучного інтелекту (ШІ) дозволяє частково зняти це навантаження, оскільки такі системи можуть виконувати попередню обробку та інтерпретацію даних, формувати попередні висновки та рекомендації для лікаря.

Розробка систем для діагностики захворювань за допомогою ШІ відкриває нові можливості як для лікарів, так і для пацієнтів. Серед таких можливостей варто відзначити віддалений моніторинг стану здоров'я, автоматичне виявлення відхилень у показниках лабораторних аналізів, швидку інтерпретацію медичних зображень, таких як рентгенограми та ЕКГ [1]. Це особливо важливо в умовах розвитку телемедицини та дистанційного надання медичних послуг, що набуває все більшого значення як в Україні, так і у світі.

Ще одним важливим аспектом є зростаюча роль великих даних (Big Data) у медицині. Щороку накопичується величезна кількість клінічної інформації, яка містить цінні закономірності для діагностики та прогнозування перебігу захворювань. Традиційні підходи до аналізу таких даних часто виявляються недостатньо ефективними, тоді як використання алгоритмів машинного навчання дозволяє виявляти приховані залежності й створювати більш точні прогностичні моделі [2]. Це відкриває нові горизонти для розвитку персоналізованої медицини.

Системи штучного інтелекту можуть суттєво сприяти розвитку доказової медицини. Вони здатні інтегрувати дані численних досліджень та клінічних спостережень, надаючи лікарям можливість приймати рішення на основі перевірених даних, а не лише власного досвіду або думки окремих фахівців. Це дозволяє зменшити ймовірність суб'єктивних помилок та підвищити якість надання медичної допомоги.

Дослідження у сфері застосування глибокого навчання для аналізу медичних даних вже продемонстрували свою високу ефективність. Наприклад, відомі моделі для інтерпретації рентгенограм легень здатні досягати точності, порівняної з результатами роботи лікарів-радіологів[3]. У сфері аналізу ЕКГ також досягнуто значного прогресу завдяки використанню глибоких нейронних мереж, що дозволяють підвищити точність виявлення серцевих патологій [4].

Розробка інтелектуальної системи для діагностики захворювань на основі медичних даних у контексті цієї дипломної роботи спрямована не лише на створення інструменту для підтримки лікарів у повсякденній практиці, але й на внесення внеску у розвиток медичних досліджень, оскільки така система дозволяє аналізувати великі масиви клінічних даних і виявляти нові закономірності. Це сприяє розвитку персоналізованої медицини та підвищенню якості медичних послуг як в Україні, так і в глобальному масштабі.

## 1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

### 1.1 Історія розвитку технологій медичної діагностики за допомогою штучного інтелекту

Використання штучного інтелекту (ШІ) у сфері медицини має багаторічну історію, яка бере свій початок ще з 70-х років ХХ століття. Перші спроби застосування інтелектуальних систем у медичній діагностиці були пов'язані з розробкою експертних систем – програмних рішень, що імітували логіку мислення спеціалістів.

Однією з перших відомих систем стала MYCIN (1972, Стенфордський університет), яка була створена для діагностики бактеріальних інфекцій крові та вибору антибіотикотерапії. Система використовувала правила типу «if-then» і базу знань, створену експертами. Хоча вона показувала високу точність, MYCIN так і не була впроваджена в клінічну практику через складність юридичного регулювання та недовіру з боку медичної спільноти.

У 80-х роках було створено кілька інших систем, зокрема INTERNIST-I та CADUCEUS, орієнтованих на діагностику внутрішніх хвороб. Проте всі ці системи мали суттєві обмеження: вони вимагали ручного наповнення баз знань, були негнучкими та не могли адаптуватися до нових типів даних.

Початок ХХІ століття ознаменувався вибуховим розвитком методів машинного навчання (ML) та, згодом, глибокого навчання (Deep Learning, DL). Це стало можливим завдяки зростанню обчислювальних потужностей, появі великих обсягів даних (Big Data) та вдосконаленню математичних алгоритмів. Замість ручного кодування знань, системи навчалися самостійно на великих масивах медичних даних: зображеннях, електрокардіограмах, результатах лабораторних аналізів тощо.

У 2016 році компанія Google DeepMind представила систему DeepMind Health, яка з високою точністю виявляла захворювання очей за

знімками сітківки. Вона продемонструвала, що глибокі нейронні мережі можуть досягати або перевищувати рівень діагностики досвідчених лікарів.

Ще одним прикладом є IBM Watson Health – потужна система, яка аналізує великі обсяги медичної інформації (наукові статті, анамнез, протоколи) і допомагає лікарям у прийнятті клінічних рішень, зокрема у сфері онкології.

Сьогодні інтелектуальні системи застосовуються у різних галузях медицини: від радіології та дерматології до кардіології, неврології і навіть психіатрії. Вони використовуються не лише для діагностики, але й для прогнозування розвитку захворювань, оцінки ризиків, персоналізованого підбору лікування та опрацювання електронної медичної документації.

Таким чином, історія розвитку ШІ в медицині – це перехід від статичних, обмежених експертних систем до адаптивних, самонавчальних моделей, які здатні інтегруватися в сучасну клінічну практику і підтримувати лікаря на всіх етапах прийняття рішень.

## 1.2 Типи медичної діагностики та види даних, які використовуються в інтелектуальних системах

Медична діагностика – це процес виявлення захворювання або патологічного стану шляхом збору, аналізу та інтерпретації клінічних даних. У сучасній клінічній практиці застосовується низка підходів до діагностики, які можна умовно класифікувати на фізикальні, лабораторні, інструментальні, функціональні та клініко-інформаційні. Кожен з них формує специфічний тип вхідних даних для систем штучного інтелекту.

### 1.2.1 Фізикальні методи діагностики

Фізикальні методи становлять базову ланку клінічного обстеження пацієнта й застосовуються лікарями на етапі первинного огляду. Ці методи

мають давню історію та базуються на безпосередньому контакті лікаря з пацієнтом, що дозволяє зібрати первинну інформацію про стан здоров'я.

Основні методи фізикального обстеження:

– огляд. Полягає у візуальній оцінці зовнішнього вигляду пацієнта: стан шкіри, колір слизових оболонок, форма грудної клітки, наявність видимих деформацій, висипів, набряків, рухова активність, вираз обличчя тощо. Наприклад, жовтяничність шкіри може вказувати на гепатобіліарні розлади;

– пальпація. Метод ручного обмацування тканин та органів. Дозволяє оцінити щільність, розміри, наявність болючості або патологічних утворень. Застосовується, зокрема, при обстеженні живота, лімфатичних вузлів, щитоподібної залози;

– перкусія. Простукування певних ділянок тіла з метою визначення фізичних властивостей підлеглих структур (резонанс, притуплення, тимпаніт). Дозволяє оцінити межі органів (наприклад, печінки, легенів), наявність рідини чи повітря;

– аускультация. Вислуховування звуків, що виникають при роботі органів (легені, серце, кишківник) за допомогою фонендоскопа. Може виявити хрипи в легенях, шуми серця, знижені або посилені перистальтичні шуми в животі.

Фізикальні методи залишаються основою клінічного мислення, проте їх ефективність значно підвищується за рахунок доповнення об'єктивними цифровими даними. Інтеграція штучного інтелекту на цьому рівні ще перебуває в стадії розвитку.

### 1.2.2 Лабораторна діагностика

Лабораторна діагностика є одним із ключових джерел об'єктивної клінічної інформації, яка дозволяє оцінити функціональний стан органів і систем пацієнта. Дані лабораторних досліджень широко використовуються

як у рутинній практиці, так і в складних діагностичних випадках, оскільки вони стандартизовані, кількісні та легко піддаються формалізації.

Ці дані, зазвичай представлені у табличному форматі, є ідеальним вхідним матеріалом для систем машинного навчання (ML) та глибокого навчання (DL), оскільки мають чітку структуру, відносно невисоку варіативність і велику кількість накопичених прикладів.

Основні типи лабораторних досліджень.

Загальний аналіз крові (ЗАК). Містить такі показники, як рівень гемоглобіну, еритроцити, лейкоцити з лейкоцитарною формулою, тромбоцити, ШОЕ тощо.

Значення. Застосовується в розпізнаванні анемії, запалення, інфекційних процесів, онкогематологічних захворювань.

Біохімічний аналіз крові. Оцінює метаболізм: глюкоза, сечовина, креатинін, електроліти, білірубін, ферменти (АЛТ, АСТ), холестерин і його фракції тощо.

Значення. Важливий у діагностиці цукрового діабету, ниркової, печінкової, серцевої недостатності.

Гормональні аналізи. Наприклад, ТТГ, Т3, Т4, кортизол, інсулін, естрогени, тестостерон.

Значення. Дають змогу оцінити функцію ендокринної системи.

Імунологічні та серологічні дослідження. Наприклад: ревматоїдний фактор, С-реактивний білок, антистрептолізин-О, ANA, антитіла до вірусів.

Значення. Використовуються при аутоімунних, інфекційних та ревматологічних хворобах.

Онкомаркери. Наприклад, PSA, СЕА, СА-125, СА-19-9.

Значення. Сприяють ранньому виявленню пухлинних процесів, але мають обмежену специфічність.

Аналіз сечі та інші біологічні рідини. Включає фізико-хімічні властивості, мікроскопію осаду, виявлення білка, глюкози, кетонових тіл.

Значення. Важливі в нефрології, урології, ендокринології.

У контексті розробки систем штучного інтелекту, лабораторні показники найчастіше представлені у вигляді табличних даних. Це дозволяє застосовувати класичні алгоритми машинного навчання – логістичну регресію, дерева рішень, Random Forest, XGBoost, нейронні мережі тощо.

Одним з класичних прикладів застосування ШІ до лабораторної діагностики є побудова моделі для визначення ризику цукрового діабету 2 типу. Зокрема, у відкритому наборі PIMA Indians Diabetes Dataset, який активно використовується в наукових дослідженнях, містяться такі клініко-лабораторні параметри: кількість вагітностей, рівень глюкози в крові, артеріальний тиск, товщина шкірної складки, рівень інсуліну, індекс маси тіла (ІМТ), віковий фактор. Моделі на цих даних досягають точності 75-85% при класифікації пацієнтів за наявністю або відсутністю діабету.

Таким чином, лабораторна діагностика виступає одним з найбільш надійних і зручних джерел даних для побудови інтелектуальних діагностичних систем. Її точність, доступність і стандартизованість відкривають широкі можливості для автоматизації прийняття рішень у медичній практиці.

### 1.2.3 Інструментальна діагностика

Інструментальна діагностика – це сукупність методів візуалізації та апаратного контролю, які дають змогу безпосередньо оцінювати морфологічний і функціональний стан внутрішніх органів та тканин. У клінічній практиці такі методи є незамінними для виявлення патологій, підтвердження діагнозів і контролю динаміки захворювання.

На відміну від лабораторних або фізикальних даних, результати інструментальних досліджень зазвичай представлені у вигляді цифрових зображень або графічних сигналів, що вимагає принципово інших методів аналізу, зокрема глибокого навчання.

Основні методи інструментальної діагностики.

Рентгенографія. Найпоширеніший метод, що дозволяє оцінити стан кісток, легенів, грудної клітки.

Значення. Часто використовується для діагностики пневмонії, туберкульозу, переломів кісток.

Комп'ютерна томографія (КТ). Формує послідовність пошарових зображень, які можна об'єднати у 3D-модель.

Значення. Дає змогу точно локалізувати пухлини, вогнища крововиливу, тромбоемболії.

Магнітно-резонансна томографія (МРТ). Забезпечує високу роздільну здатність м'яких тканин, особливо ефективна в нейро- та ортопедичній діагностиці.

Значення. Застосовується для виявлення пухлин мозку, склерозу, гриж міжхребцевих дисків.

Ультразвукове дослідження (УЗД). Універсальний метод оцінки органів черевної порожнини, серця (ехокардіографія), нирок, щитоподібної залози, плоду.

Значення. Важливий у вагітності, ендокринології, кардіології.

Мамографія, ангіографія, ендоскопія. Більш специфічні методи, що мають вузьке клінічне застосування, але також генерують візуальні дані для аналізу.

Для обробки інструментальних зображень найчастіше застосовуються сверточні нейронні мережі (CNN), які здатні виділяти ознаки з вхідних зображень без попереднього ручного втручання.

Одним із найвідоміших прикладів успішного застосування глибокого навчання в медичній візуальній діагностиці є модель CheXNet, розроблена дослідниками зі Стенфордського університету в 2017 році.

Метою проекту було створення нейронної мережі, здатної автоматично виявляти пневмонію на рентгенограмах грудної клітки. Для цього було використано публічний датасет ChestX-ray14 від NIH, який містить понад 100 000 знімків з 14 типами патологій.

CheXNet була побудована на основі глибокої мережі DenseNet-121 та навчена з використанням технологій глибокого навчання для багатокласової класифікації. Після навчання модель змогла досягти точності, яка перевищувала середній рівень радіологів у завданні виявлення пневмонії.

Цей проєкт став важливою віхою у визнанні потенціалу ШІ у медичній візуалізації, а також показав, що за наявності великого, якісно анотованого датасету можливо створити моделі, які не лише підтримують, але й у деяких випадках перевершують фахівців у вузькоспеціалізованих завданнях.

#### 1.2.4 Функціональна діагностика

Функціональна діагностика – це сукупність методів, які дозволяють оцінити динамічні процеси життєдіяльності організму шляхом реєстрації електричних, звукових або інших фізіологічних сигналів. Ці дослідження є важливими для виявлення порушень ритму, функцій органів, рівня насичення киснем тощо.

Особливістю даних функціональної діагностики є те, що вони часто подаються у вигляді часових рядів або сигналів, що вимагає використання спеціалізованих методів аналізу з боку ШІ – таких як рекурентні нейронні мережі (RNN), LSTM, а також 1D-CNN або трансформери для сигналів.

Основні методи функціональної діагностики.

Електрокардіографія (ЕКГ). Реєстрація електричної активності серця. Значення. Застосовується для виявлення аритмій, ішемії, гіпертрофії, блокад.

Електроенцефалографія (ЕЕГ). Запис електричної активності мозку. Значення. Важлива для діагностики епілепсії, сну, енцефалопатій, коматозних станів.

Спірографія. Оцінка об'єму і швидкості дихання.

Значення. Діагностика бронхіальної астми, ХОЗЛ, рестриктивних порушень.

Пульсоксиметрія. Визначення насичення крові киснем (SpO<sub>2</sub>).

Значення. Швидка оцінка респіраторної функції, особливо у пацієнтів із COVID-19.

Відомий датасет MIT-BIH Arrhythmia Dataset з PhysioNet містить розмічені 30-хвилинні ЕКГ-записи, які активно використовуються для навчання моделей ШІ. У наукових роботах на цьому наборі застосовують 1D-CNN або LSTM, які розпізнають тип аритмії (фібриляція передсердь, блокади, екстрасистолії тощо) з точністю понад 90%.

Функціональна діагностика є важливою галуззю, де штучний інтелект дозволяє автоматизувати інтерпретацію складних сигналів та виявляти відхилення ще до появи клінічних симптомів. Системи, здатні аналізувати ЕКГ, ЕЕГ та інші сигнали в реальному часі, відкривають нові горизонти у персоналізованій медицині, телеметрії та віддаленому моніторингу пацієнтів.

### 1.3 Актуальність задачі медичної діагностики методами штучного інтелекту

Охорона здоров'я є однією з найкритичніших і найскладніших галузей людської діяльності, де від точності й своєчасності діагностики залежить не лише якість життя, а й його збереження. У той самий час сучасна медицина стикається з низкою викликів, які ускладнюють процес діагностування: зростання кількості пацієнтів та навантаження на лікарів, дефіцит кваліфікованих фахівців, особливо в сільських та віддалених регіонах, високий рівень діагностичних помилок, зумовлений людським фактором, втому, нестачею досвіду, складність інтерпретації великої кількості даних, які генеруються при обстеженнях.

Застосування методів ШІ у медицині дозволяє не лише автоматизувати процеси аналізу великих обсягів даних, але й істотно підвищити точність діагнозів, стандартизувати клінічне мислення та знизити кількість людських помилок. За даними Всесвітньої організації охорони здоров'я, діагностичні помилки трапляються в 10–15% клінічних випадків, що суттєво впливає на рівень смертності та ускладнень. Сучасні системи ШІ, навчені на великих обсягах клінічної інформації, здатні забезпечити підтримку у формулюванні первинного діагнозу, виявляти патології на зображеннях, аналізувати біомедичні сигнали, інтерпретувати текстові записи в електронній медичній документації.

Окрему актуальність тема набуває з огляду на зростання обсягу медичних даних. За оцінками аналітиків, понад 80% усіх клінічних даних є неструктурованими – це зображення, сигнали, текстові виписки, які складно обробляти класичними методами. Саме штучний інтелект здатен «перетворити» ці дані на цінну інформацію для діагностики, забезпечуючи автоматизовану інтерпретацію зображень (рентген, КТ, МРТ), аналіз електрокардіограм, визначення ризику захворювань за аналізами крові, виявлення симптомів у скаргах пацієнтів.

Водночас важливо відзначити, що вже сьогодні ШІ демонструє реальні результати у клінічній практиці. Так, системи комп'ютерного зору виявляють патології на знімках із точністю, що не поступається рівню досвідчених лікарів-радіологів. У галузі кардіології штучні нейронні мережі класифікують порушення ритму серця за ЕКГ з точністю понад 90%. У текстовій діагностиці трансформерні моделі, такі як ClinicalBERT, використовуються для автоматичного аналізу історій хвороби та витягання клінічно значущих відомостей.

Не менш важливою є й соціально-економічна значущість впровадження таких рішень. Підвищення точності діагностики знижує кількість повторних госпіталізацій, скорочує витрати системи охорони здоров'я та дозволяє більш раціонально використовувати кадрові ресурси.

Такі технології особливо ефективні в умовах телемедицини, мобільної медицини та під час роботи в надзвичайних ситуаціях, зокрема під час пандемій або воєнного часу.

Таким чином, задача діагностики захворювань методами штучного інтелекту є не лише науково обґрунтованою, а й практично необхідною. Вона відповідає сучасним тенденціям цифрової медицини та медичної інформатики, створюючи потенціал для покращення якості медичних послуг, зниження ризику помилок і підвищення доступності охорони здоров'я. Усе це обумовлює високу актуальність обраної теми дослідження як на національному, так і на глобальному рівні.

#### 1.4 Мета роботи та постановка задачі

##### 1.4.1 Мета роботи

Метою даної кваліфікаційної роботи є дослідження, розробка та апробація інтелектуальної системи підтримки медичної діагностики із застосуванням сучасних методів штучного інтелекту. У рамках поставленої мети передбачається створення програмного прототипу, здатного на основі вхідних клінічних даних – таких як лабораторні показники, фізіологічні параметри чи медичні зображення – виконувати автоматизований аналіз, виявлення аномалій та класифікацію ймовірного захворювання.

Особливу увагу в роботі приділено вибору оптимальних алгоритмів машинного навчання або глибокого навчання відповідно до типу вхідних даних, побудові ефективної моделі з високими показниками точності та інтерпретованості результатів, а також оцінці її придатності для використання в реальних умовах клінічної практики або як частини телемедичних сервісів.

Таким чином, мета дослідження охоплює як теоретичне вивчення можливостей застосування штучного інтелекту в медичній діагностиці, так

і реалізацію прикладної моделі з практичною цінністю в умовах цифрової трансформації системи охорони здоров'я.

#### 1.4.2 Постановка задачі

Для досягнення поставленої мети необхідно вирішити низку взаємопов'язаних наукових, технічних та прикладних задач, зокрема:

- провести аналіз сучасного стану застосування методів штучного інтелекту в медичній діагностиці, визначити основні підходи, що демонструють високу ефективність при роботі з різними типами медичних даних;

- класифікувати типи вхідної медичної інформації (зображення, лабораторні аналізи, сигнали, текстові документи) та проаналізувати відповідність кожного з них певним категоріям алгоритмів штучного інтелекту;

- визначити критерії вибору моделей машинного або глибокого навчання, що найкраще підходять до обраної задачі діагностики, та обґрунтувати доцільність їх застосування;

- реалізувати програмний прототип інтелектуальної системи на основі обраних моделей, провести навчання та валідацію з використанням сучасних фреймворків;

- провести оцінювання якості моделі за метриками точності, повноти, F1-міри, а також проаналізувати її стабільність і потенціал клінічного застосування;

- підготувати висновки за результатами реалізації, оцінити доцільність використання запропонованої системи та сформулювати рекомендації щодо подальшого розвитку та впровадження.

## 2 ОГЛЯД ТА ПОРІВНЯЛЬНИЙ АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА ТЕХНОЛОГІЙ

### 2.1 Сучасний стан розробок у галузі медичних AI-систем

Протягом останніх двох десятиліть медична галузь зазнає стрімкої трансформації під впливом цифрових технологій, зокрема – методів штучного інтелекту. Якщо на початку 2000-х років такі системи переважно мали статус експериментальних або дослідницьких розробок, то сьогодні багато з них вже інтегруються у практику охорони здоров'я, проходять сертифікацію, включаються до телемедичних платформ або застосовуються у форматі мобільних застосунків.

Згідно з даними аналітичних звітів McKinsey, Markets&Markets та Deloitte, глобальний ринок рішень на основі ШІ в медицині зростає щорічно на понад 35% і, за прогнозами, сягне понад 100 млрд доларів США до 2030 року. Це свідчить про не лише наукову та технологічну зацікавленість, а й про практичну потребу у впровадженні таких систем у реальні клінічні процеси.

Найбільш активно ШІ застосовується в тих галузях медицини, де спостерігається великий обсяг даних, висока стандартизація обстежень і можливість цифрової інтерпретації результатів. Серед таких напрямів – радіологія, офтальмологія, дерматологія, кардіологія, лабораторна діагностика та обробка електронної медичної документації.

Однією з ключових рушійних сил розвитку ІІ в медицині стало накопичення великих масивів відкритих даних (open medical datasets), розвиток хмарних обчислень та доступність фреймворків для машинного і глибокого навчання. Такі сервіси, як TensorFlow, PyTorch, HuggingFace Transformers, Google Cloud Healthcare API значно спростили процес створення та тестування діагностичних моделей навіть у рамках невеликих дослідницьких груп.

Значний прогрес спостерігається й у напрямі пояснювального штучного інтелекту (Explainable AI), що дозволяє підвищити довіру до результатів моделей – особливо в медичному середовищі, де прозорість прийняття рішень є критично важливою. Відповідно, у сучасних системах часто застосовуються інтерактивні теплові карти, метрики значущості ознак, а також механізми перевірки внутрішньої логіки моделі.

Крім того, спостерігається зростання кількості клінічних досліджень, у яких ШІ застосовується як додатковий інструмент до традиційних методів діагностики. У багатьох країнах вже створено нормативні бази, що дозволяють сертифікувати медичні AI-системи. Так, Управління з контролю за продуктами та ліками США (FDA) вже затвердило понад 500 AI-рішень для використання у клінічній практиці, зокрема – у діагностиці пневмонії, аналізі мамограм, скринінгу діабетичної ретинопатії тощо.

Таким чином, станом на сьогодні можна говорити про перехід систем ШІ у медицині з розряду «інноваційних концепцій» у повноцінні інструменти для підтримки прийняття рішень у клінічному середовищі. Наступні підпункти будуть присвячені аналізу найбільш відомих і ефективних реалізацій таких систем, а також порівнянню технологічних підходів, які лежать в їх основі.

## 2.2 Огляд існуючих програмних рішень

### 2.2.1 CheXNet

Модель CheXNet була представлена дослідниками зі Стенфордського університету у 2017 році як один із перших вдалих прикладів використання глибоких згорткових нейронних мереж (CNN) для автоматичної інтерпретації рентгенографічних знімків грудної клітки. Основною метою проєкту було створення системи, здатної виявляти пневмонію з точністю, не нижчою за рівень практикуючих радіологів.

Основою для навчання слугували дані з ChestX-ray14 – великого відкритого датасету, опублікованого Національним інститутом охорони здоров'я США (NIH) [5]. Датасет містить понад 112 тисяч рентген-знімків від приблизно 30 тисяч пацієнтів, анотації до яких охоплюють 14 типів легеневих патологій, зокрема пневмонію, ателектаз, кардіомегалію, плевральний випіт тощо. Важливо, що знімки були зібрані з реальної клінічної практики, що забезпечує різноманітність за якістю, ракурсами та ступенем вираженості патологій – фактори, які наближають модель до реального середовища використання.

CheXNet побудована на основі архітектури DenseNet-121 – глибокої CNN, що забезпечує ефективне поширення ознак між шарами та високу точність навіть при обмеженому обсязі міток. Модель була попередньо навчена на ImageNet (для загального розпізнавання зображень), а потім донавчена на ChestX-ray14, що дало змогу використати підхід Transfer Learning – тобто перенесення вже «вивчених» загальних ознак до конкретного медичного завдання.

Після навчання модель пройшла оцінювання на незалежній вибірці, де її продуктивність порівнювали з результатами чотирьох досвідчених радіологів. За метрикою F1-score, CheXNet показала кращий або рівнозначний результат у виявленні пневмонії, що стало важливим аргументом на користь потенціалу глибоких моделей у радіологічній діагностиці.

Крім точності, дослідники також забезпечили можливість інтерпретації рішень моделі. Для цього використовувався метод Class Activation Mapping (CAM), який дає змогу візуалізувати ділянки знімка, що найбільше вплинули на прогноз (рисунок 2.1). Це підвищує довіру лікарів до автоматизованої системи і дозволяє використовувати її як інструмент підтримки рішень, а не як «чорну скриньку».



Рисунок 2.1 – Теплова карта, що локалізує області зображення, найбільш характерні для пневмонії

CheXNet відкрив нову епоху в застосуванні ШІ для обробки медичних зображень, продемонструвавши, що моделі на основі CNN можуть не лише відтворювати, а й перевершувати деякі аспекти людської діагностичної здатності. Цей приклад також став основою для десятків подальших досліджень і комерційних розробок у галузі рентгенології та комп'ютерного зору в медицині.

### 2.2.2 DeepMind Health

Компанія DeepMind Health, дочірній підрозділ Google, з 2016 року веде активні дослідження у галузі застосування штучного інтелекту в медицині, зокрема – в офтальмології. Одним із найвідоміших її проєктів стало створення глибокої нейронної мережі для аналізу оптичної когерентної томографії сітківки (ОКТ) з метою виявлення серйозних захворювань очей, зокрема вікової макулярної дегенерації, діабетичної ретинопатії, глаукоми та набряку сітківки.

Проєкт реалізовувався у співпраці з Мурфілдською очною лікарнею (Moorfields Eye Hospital) у Лондоні. Для навчання моделі було

використано близько 15 тисяч сканів сітківки, які були розмічені висококваліфікованими офтальмологами. Це забезпечило високу якість анотацій, що є ключовим чинником для ефективного навчання моделей глибокого навчання.

На відміну від багатьох попередніх підходів, у цьому проєкті використовувалась двоступенева архітектура:

- сегментаційна мережа, яка аналізувала сирі ОКТ-скани та будувала структуровану карту сітківки – з виявленням анатомічних зон і патологічних змін;

- класифікаційна модель, яка на основі цієї карти робила висновок про наявність захворювання та необхідність термінового направлення до лікаря.

Це дозволило зберегти медичну інтерпретованість на всіх етапах, лікар не просто отримував діагноз, а бачив, які саме анатомічні зміни привели модель до відповідного висновку.

Після навчання система досягла точності на рівні або вищої, ніж у восьми досвідчених офтальмологів, зокрема в задачі класифікації випадків, які потребують негайного втручання. Згідно з публікацією у журналі *Nature Medicine* (2018 р.), модель DeepMind досягла 94% точності у розпізнаванні станів, що потребують термінового лікування, з високою чутливістю та специфічністю.

Окрім цього, модель продемонструвала здатність узагальнювати знання – вона ефективно працювала не лише на сканах, на яких навчалась, а й на даних з інших клінік і пристроїв, що вказує на добру переносимість.

Ця система не тільки допомагає лікарям у прийнятті рішень, але й має потенціал стати основою для масових офтальмологічних скринінгів, особливо у регіонах з обмеженим доступом до фахівців.

### 2.2.3 Watson for Oncology

Watson for Oncology – це система штучного інтелекту, розроблена компанією IBM у партнерстві з Memorial Sloan Kettering Cancer Center (США), яка призначена для підтримки прийняття клінічних рішень у сфері онкології. На відміну від попередніх прикладів (CheXNet, DeepMind), які орієнтовані на аналіз зображень або сигналів, Watson працює з текстовими, семіструктурованими та неструктурованими медичними даними – історіями хвороби, результатами обстежень, протоколами лікування.

Watson for Oncology аналізує електронну медичну документацію пацієнта, поєднуючи її з величезним масивом наукових публікацій, клінічних рекомендацій, баз даних лікарських препаратів, інформацією з онкологічних протоколів, а також з власною базою знань, яка постійно оновлюється. У результаті система пропонує лікарю ранжований перелік варіантів лікування, класифікованих за рівнем впевненості й обґрунтованих науковими джерелами.

Архітектура Watson включає компоненти Natural Language Processing (NLP), Rule-based reasoning та Machine Learning. Для інтерпретації текстів використовується спеціалізований модуль обробки природної мови, натренований на медичній термінології. Цей модуль вміє розпізнавати клінічні терміни, виявляти причинно-наслідкові зв'язки між симптомами, діагнозами, біомаркерами та схемами лікування.

Система проходила тестування в кількох онкоцентрах у США, Індії, Китаї, Південній Кореї, Бразилії. У ряді досліджень було показано, що вибір Watson збігався з думкою комісії онкологів у 70-90% випадків, залежно від нозології та рівня локалізації раку. Особливо успішними виявилися сценарії, де пацієнт мав рідкісну форму пухлини або супутні захворювання, що ускладнювали типову схему лікування.

Однак, попри технологічні досягнення, Watson for Oncology не є автономною системою, а функціонує виключно як інструмент підтримки, залишаючи остаточне рішення за лікарем. Серед викликів, з якими зіткнулась система, складність адаптації до локальних клінічних протоколів, мовних бар'єрів, а також необхідність ручного налаштування для кожного медичного закладу.

Втім, цей приклад залишається еталоном інтеграції ШІ в складну, багатофакторну галузь – онкологічну допомогу, де рішення мають враховувати десятки змінних, медичних норм, супутніх станів і показників ефективності.

### 2.3 Методи, що використовуються в існуючих програмних рішеннях

Існуючі AI-системи, що застосовуються в медичній діагностиці, ґрунтуються на широкому спектрі алгоритмів машинного та глибокого навчання. Конкретний вибір методів визначається типом вхідних даних (зображення, текст, сигнали, структуровані таблиці), характером задачі (класифікація, сегментація, прогнозування) і вимогами до точності, інтерпретованості, швидкодії.

#### 2.3.1 Глибоке навчання для медичних зображень

Глибоке навчання (Deep Learning, DL) є сучасним підходом штучного інтелекту, що знайшов широке застосування у медичній візуалізації та аналізі зображень. Ця технологія ґрунтується на багаторівневих штучних нейронних мережах, здатних автоматично виділяти ознаки зображень та встановлювати складні закономірності в даних без потреби в ручному створенні правил або шаблонів. Розвиток глибокого навчання відкрив нові горизонти у медичній діагностиці, дозволяючи підвищити точність

виявлення патологій, пришвидшити обробку обстежень та стандартизувати підхід до інтерпретації медичних даних.

Однією з найбільш поширених архітектур глибокого навчання для роботи з медичними зображеннями є згорткові нейронні мережі (Convolutional Neural Networks, CNN). Вони забезпечують ефективну обробку двовимірних та тривимірних зображень, навчаючись автоматично виділяти локальні та глобальні ознаки – від простих геометричних контурів до складних морфологічних структур. Значного поширення набула архітектура U-Net, яка спеціально розроблена для завдань сегментації медичних зображень і дозволяє досягати високої точності при виділенні контурів пухлин, органів чи патологічних вогнищ. Інші модифікації CNN, такі як ResNet, DenseNet та EfficientNet, використовуються для підвищення глибини та стабільності моделей за рахунок залишкових зв'язків і покращеної передачі градієнтів під час навчання. Новітні розробки включають використання Vision Transformers (ViT), які демонструють конкурентоспроможність або навіть перевагу порівняно з класичними CNN завдяки глобальному підходу до аналізу зображень і механізмам уваги [6].

Глибоке навчання вже активно використовується у низці клінічних напрямів. У рентгенології моделі глибокого навчання довели свою ефективність у розпізнаванні патологій на рентгенограмах грудної клітки, таких як пневмонія, плевральний випіт та пневмоторакс. Відомий приклад – модель CheXNet, розроблена Стенфордським університетом, яка показала точність на рівні або вище середнього рівня лікарів-рентгенологів. Комерційні рішення, як-от Lunit INSIGHT CXR, вже впроваджені у клінічну практику в десятках країн світу, забезпечуючи автоматичну інтерпретацію знімків та підвищуючи продуктивність роботи медичних закладів. У сфері маммографії глибоке навчання демонструє значний потенціал для раннього виявлення раку молочної залози. Модель Google Health AI for Breast Cancer

Screening у дослідженнях, опублікованих у журналі Nature, перевершила лікарів у ряді показників, зокрема в точності розпізнавання злоякісних змін.

Магнітно-резонансна томографія та комп'ютерна томографія також є важливими областями застосування глибокого навчання. Архітектура U-Net використовується для сегментації пухлин мозку, наприклад, у рамках міжнародного змагання BraTS Challenge, що стимулює розвиток інновацій у цій сфері. Платформи як Aidoc забезпечують аналіз КТ-зображень в реальному часі для виявлення ішемічного інсульту або внутрішньочерепних крововиливів і вже інтегровані в клінічну практику багатьох медичних центрів. У дерматології моделі глибокого навчання застосовуються для розпізнавання новоутворень шкіри, зокрема меланоми. Змагання ISIC Challenge сприяли створенню рішень, здатних демонструвати продуктивність на рівні експертів-дерматологів [7]. В офтальмології FDA вперше схвалила автономну систему IDx-DR на базі глибокого навчання для виявлення діабетичної ретинопатії без участі лікаря на етапі скринінгу.

Принцип роботи систем глибокого навчання передбачає кілька ключових етапів. На першому етапі здійснюється підготовка даних, яка включає нормалізацію, аугментацію для збільшення різноманітності тренувальних прикладів, видалення артефактів та шумів. Далі модель навчається на великій кількості зображень, поступово налаштовуючи свої параметри для мінімізації помилок при прогнозуванні. Процес навчання часто поєднується з валідацією на окремих підмножинах даних, щоб запобігти перенавчанню та оцінити загальну здатність моделі до узагальнення. Після завершення навчання модель тестується на нових даних для перевірки її ефективності в умовах, наближених до реальних клінічних сценаріїв. Кінцевим етапом є застосування моделі у клінічній практиці, коли вона використовується для аналізу нових медичних зображень у режимі реального часу або у вигляді допоміжного інструменту для лікарів.

Попри численні досягнення, глибоке навчання в медичній візуалізації стикається з певними викликами. Однією з основних проблем є обмежений

доступ до великих та якісно анотованих медичних датасетів через питання конфіденційності пацієнтів та етичні обмеження. Іншою важливою проблемою залишається пояснюваність рішень моделей глибокого навчання, оскільки їхня структура часто сприймається як «чорна скринька», що ускладнює інтерпретацію результатів лікарями та пацієнтами. Водночас перспективи розвитку цієї технології є надзвичайно широкими. Передбачається подальше вдосконалення мультимодальних моделей, які будуть поєднувати дані з медичних зображень, електронних медичних записів, геноміки та інших джерел для створення більш точних та персоналізованих систем діагностики. Зокрема, розвиток мобільних додатків і телемедичних платформ із вбудованими алгоритмами глибокого навчання сприятиме підвищенню доступності якісної медичної допомоги у віддалених та малозабезпечених регіонах.

### 2.3.2 Машинне навчання

Машинне навчання відіграє значну роль у розвитку медичних інформаційних систем, діагностичних засобів і систем підтримки прийняття клінічних рішень. На відміну від традиційних алгоритмів, що базуються на чітко визначених правилах, машинне навчання дає змогу будувати моделі, які самостійно виявляють закономірності в даних та роблять прогнозування або класифікацію на основі накопиченого досвіду. Це відкриває нові можливості для обробки великих обсягів різноманітної медичної інформації – від лабораторних аналізів і фізіологічних параметрів до даних електронних медичних записів та результатів візуалізаційних досліджень.

У сучасній медичній практиці застосовуються різноманітні підходи машинного навчання, серед яких найбільш поширені моделі на основі методів класифікації, регресії та кластеризації. Серед них особливе місце займають дерева рішень, випадкові ліси (Random Forest), метод опорних векторів (Support Vector Machine, SVM), градієнтний бустинг (наприклад,

XGBoost, LightGBM) та ансамблеві методи. Ці алгоритми дозволяють створювати моделі, здатні прогнозувати наявність захворювань (наприклад, діабету, ішемічної хвороби серця, серцевої недостатності), оцінювати ризики ускладнень або смертності, виявляти групи пацієнтів зі схожими клінічними характеристиками. При цьому моделі можуть бути як лінійними, що забезпечують простоту інтерпретації, так і нелінійними, що дозволяють вловлювати складні взаємозв'язки між змінними [8].

Особливу роль у застосуванні машинного навчання в медицині відіграє попередня обробка даних та підбір ознак (feature selection), оскільки якість вхідної інформації значною мірою визначає точність побудованої моделі. Медичні дані часто характеризуються наявністю пропусків, аномальних значень або сильно корельованих ознак, що потребує спеціальних підходів до їх очищення й трансформації. Для цього застосовуються методи нормалізації, масштабування, імпутації відсутніх значень, а також алгоритмічні підходи для автоматичного відбору найбільш інформативних ознак. Значного поширення набули також техніки балансування класів, як-от oversampling і undersampling, особливо актуальні для завдань виявлення рідкісних захворювань або ускладнень [9].

Машинне навчання знайшло практичне застосування в різних галузях медицини. Зокрема, алгоритми на основі Random Forest та XGBoost активно використовуються для прогнозування ризику розвитку серцево-судинних захворювань за даними традиційних факторів ризику, таких як артеріальний тиск, рівень холестерину, індекс маси тіла, наявність шкідливих звичок тощо. Прикладами таких рішень є моделі на основі даних Framingham Heart Study, які застосовуються в наукових дослідженнях і клінічній практиці для оцінки ймовірності розвитку ішемічної хвороби серця. У діабетології машинне навчання використовується для створення скринінгових інструментів, які на основі комбінації клініко-лабораторних показників прогнозують наявність діабету або ризик розвитку діабетичних ускладнень.

Ці моделі часто впроваджуються у вигляді модулів клінічних інформаційних систем або мобільних додатків.

У сфері онкології машинне навчання застосовується для побудови моделей, що аналізують геномні, протеомні та інші омічні дані з метою стратифікації пацієнтів, підбору терапії та прогнозування відповідей на лікування. Значного розвитку набули також моделі кластеризації, які дозволяють виявляти підтипи пухлин або інших захворювань на основі комплексних даних про пацієнтів. Такі підходи вже інтегровані в дослідницькі проєкти з персоналізованої медицини, зокрема у рамках консорціумів The Cancer Genome Atlas (TCGA) та інших міжнародних ініціатив.

Іншим перспективним напрямом є використання машинного навчання для аналізу текстових даних електронних медичних записів. Алгоритми обробки природної мови (Natural Language Processing, NLP), інтегровані в моделі машинного навчання, дозволяють автоматично виділяти ключові клінічні ознаки з неструктурованих текстів, будувати хронологію подій, виявляти взаємозв'язки між симптомами, діагнозами та призначеннями. Це відкриває можливості для створення систем підтримки прийняття рішень, здатних забезпечити лікаря релевантною інформацією у потрібний момент, а також для автоматичного формування звітів і статистики.

Перспективи застосування машинного навчання в медицині надзвичайно широкі й продовжують розширюватися завдяки розвитку методів explainable AI (пояснюваного штучного інтелекту), які дозволяють зробити роботу моделей більш прозорою та зрозумілою для лікарів і пацієнтів. Очікується, що найближчим часом машинне навчання буде все більше інтегруватися в клінічну практику у вигляді гібридних систем, які поєднують алгоритмічний аналіз даних із експертною оцінкою фахівців. Водночас розвиток цієї галузі потребує подолання низки викликів, зокрема забезпечення високої якості даних, захисту конфіденційності пацієнтів, узгодження моделей із клінічними стандартами та нормативно-правовою

базою [10]. Незважаючи на це, машинне навчання вже сьогодні змінює уявлення про можливості аналізу медичних даних та сприяє формуванню нової парадигми доказової медицини, що базується на даних великих обсягів і високої складності.

### 2.3.3 Часові ряди

Часові ряди є важливим класом даних у медицині, оскільки багато клінічних показників мають динамічний характер і змінюються з часом. До таких даних належать електрокардіограми (ЕКГ), електроенцефалограми (ЕЕГ), моніторингові дані артеріального тиску, пульсу, сатурації кисню, рівня глюкози в крові, а також дані з медичних пристроїв і сенсорів, що реєструють фізіологічні параметри пацієнтів у реальному часі. Методи аналізу часових рядів на основі машинного та глибокого навчання відіграють дедалі важливішу роль у розробці систем моніторингу здоров'я, раннього виявлення патологічних станів та прогнозування розвитку захворювань.

У роботі з часовими рядами традиційно використовуються статистичні моделі, такі як авторегресивні (AR), авторегресивно-ковзні середні (ARMA), авторегресивно-інтегровані ковзні середні (ARIMA) та моделі з сезонною компонентою (SARIMA). Ці підходи дозволяють будувати прогнози на основі історичних даних та аналізувати тренди й циклічні зміни. Проте у багатьох випадках медичні часові ряди характеризуються високою складністю, нелінійністю, значним рівнем шуму та наявністю багатьох взаємозалежних факторів. Тому все частіше для обробки таких даних застосовуються методи машинного навчання та глибоких нейронних мереж, здатні захоплювати приховані залежності й закономірності, що не піддаються опису класичними статистичними методами.

Одним із найбільш поширених підходів у цій сфері є використання рекурентних нейронних мереж (Recurrent Neural Networks, RNN) та їх модифікацій, зокрема довгокороткочасної пам'яті (Long Short-Term Memory, LSTM) і мережі з врахуванням гейтових механізмів (Gated Recurrent Unit, GRU). Ці моделі спеціально створені для роботи з послідовними даними, оскільки вони враховують залежності між елементами ряду на різних часових відстанях. У клінічній практиці LSTM широко використовуються для прогнозування ризику серцевих аритмій за даними ЕКГ, виявлення епілептичних нападів за ЕЕГ, а також для прогнозування стану пацієнтів у відділеннях інтенсивної терапії за даними безперервного моніторингу життєво важливих функцій. Наприклад, у проектах MIT Lab та PhysioNet Challenge було продемонстровано, як моделі LSTM можуть передбачати розвиток сепсису або зупинки серця задовго до клінічних проявів, що дозволяє своєчасно втручатися у перебіг патологічного процесу [11].

Окрім рекурентних мереж, значного поширення набули одномірні згорткові нейронні мережі (1D CNN), які ефективно працюють з часовими рядами завдяки здатності виділяти локальні патерни та характерні ознаки сигналу. Такі моделі успішно застосовуються для автоматичної класифікації ритмів серця, виявлення порушень дихання під час сну, розпізнавання патологічних коливань рівня глюкози та багатьох інших завдань. Нещодавно в аналізі часових рядів почали використовувати трансформери – архітектури, що демонструють високу продуктивність завдяки механізмам уваги й здатності обробляти залежності на довгих часових відрізках. Приклади таких рішень уже з'являються у проектах зі створення систем прогнозування артеріального тиску або предикції розвитку гіпоглікемічних епізодів у хворих на цукровий діабет.

Реальні приклади впровадження методів аналізу часових рядів у медицину можна знайти в роботі таких систем, як CardioNet, що використовує глибокі моделі для аналізу багатогодинних записів ЕКГ та

виявлення аритмій, та Medtronic Guardian Connect – системи безперервного моніторингу рівня глюкози з елементами прогнозування на основі машинного навчання. У відділеннях інтенсивної терапії сучасні моніторингові системи все частіше інтегрують модулі ШІ, які аналізують часові ряди даних пульсу, тиску, сатурації та температури для виявлення ранніх ознак погіршення стану пацієнтів.

Перспективи використання аналізу часових рядів у медицині тісно пов'язані з розвитком технологій персоналізованого моніторингу та телемедицини. Передбачається, що системи на основі ШІ забезпечать більш точне прогнозування загострень хронічних захворювань, підвищать ефективність домашнього спостереження за пацієнтами та дозволять впровадити проактивні підходи до лікування [12]. Водночас важливими викликами залишаються необхідність забезпечення високої якості й цілісності даних, адаптація моделей під індивідуальні особливості пацієнтів, а також вирішення етичних і правових питань, пов'язаних з обробкою медичної інформації. Розвиток методів аналізу часових рядів із використанням штучного інтелекту створює підґрунтя для формування нової парадигми медичної діагностики, де рішення базуються на безперервному потоці даних та прогнозах, що генеруються у реальному часі.

#### 2.3.4 Експертні системи

Експертні системи є одним із найстаріших і водночас фундаментальних напрямів використання штучного інтелекту в медицині. Вони призначені для моделювання процесу прийняття рішень фахівця-експерта та надання рекомендацій лікарю під час діагностики, вибору лікування чи прогнозування розвитку захворювань. Основною особливістю експертних систем є наявність бази знань – структурованого сховища

правил, фактів, гіпотез та зв'язків між ними, що відображає накопичений досвід експертів у певній галузі медицини.

Традиційно експертні системи будуються на основі продукційних правил виду «якщо – то», які описують причинно-наслідкові зв'язки між клінічними ознаками, лабораторними показниками, результатами обстежень та можливими діагнозами або терапевтичними діями. Механізм виведення системи аналізує введені користувачем дані та, використовуючи логічні правила, формує висновки. При цьому для врахування невизначеності часто використовуються методи нечіткої логіки або ймовірнісні підходи, зокрема баєсівські мережі [13]. У сучасних реалізаціях експертні системи поєднують традиційні логічні методи з елементами машинного навчання для автоматичного оновлення знань і адаптації до нових даних.

Класичними прикладами експертних систем у медицині є MYCIN – одна з перших систем для діагностики інфекційних захворювань та підбору антибіотиків, та INTERNIST-1 – система для підтримки діагностичних рішень у внутрішній медицині. У сучасній практиці експертні системи застосовуються в багатьох сферах: від автоматизованого контролю якості медичних процесів до допомоги у виборі тактики лікування онкологічних захворювань. Наприклад, у клінічній онкології активно розвиваються експертні модулі, що базуються на клінічних рекомендаціях і протоколах NCCN, ASCO та ESMO для вибору індивідуалізованої терапії пацієнтів.

Перспективи розвитку експертних систем у медицині пов'язані зі створенням гібридних платформ, де експертні правила інтегруються з результатами аналізу великих даних та алгоритмами глибокого навчання. Такі системи здатні не лише застосовувати наявні знання, а й доповнювати їх новою інформацією, отриманою з практики. Виклики, з якими стикаються розробники експертних систем, стосуються необхідності актуалізації бази знань, забезпечення сумісності з електронними медичними записами та

вирішення етичних питань, пов'язаних із відповідальністю за рекомендації системи.

### 2.3.5 Генетичні алгоритми

Генетичні алгоритми є однією з евристичних технологій штучного інтелекту, яка знайшла застосування в медицині завдяки своїй здатності ефективно вирішувати складні оптимізаційні завдання. Цей підхід базується на принципах природного відбору та еволюції в біологічних системах. Генетичні алгоритми імітують процеси мутації, кросинговеру та селекції для поступового вдосконалення рішень та пошуку глобального оптимуму в умовах багатовимірних і нелінійних просторів даних. У медицині це відкриває можливості для побудови гнучких моделей, здатних знаходити найкращі комбінації діагностичних ознак, параметрів лікування або структури прогнозуючих моделей.

Сутність роботи генетичних алгоритмів полягає у створенні популяції можливих рішень (хромосом), які представляють різні варіанти відповідей на поставлену задачу. Кожне рішення кодується у вигляді послідовності чисел або бітів і оцінюється за допомогою функції пристосованості (fitness function), що відображає його якість з погляду досягнення мети. Далі відбуваються процеси селекції найбільш пристосованих рішень, їх схрещування та випадкових мутацій для генерації нового покоління рішень. Цей цикл повторюється багаторазово, доки не буде знайдено рішення, що задовольняє задані критерії.

У медичних завданнях генетичні алгоритми використовуються для оптимізації складних діагностичних і прогностичних моделей. Наприклад, вони застосовуються для відбору найбільш інформативних ознак серед великої кількості клінічних, лабораторних і демографічних даних у завданнях прогнозування захворювань серцево-судинної системи, діабету, онкологічних патологій. Генетичні алгоритми також ефективно

використовуються при побудові ансамблевих моделей машинного навчання, коли потрібно знайти оптимальні параметри або структуру об'єднаних класифікаторів. У фармакогеноміці вони слугують для підбору комбінацій генетичних маркерів, що асоціюються з відповіддю на певні лікарські засоби або ризиком розвитку побічних ефектів.

Реальні приклади застосування генетичних алгоритмів у медицині можна знайти в дослідженнях прогнозування виживаності онкологічних пацієнтів за даними багатовимірної біомедичної інформації або в автоматизованих системах підбору параметрів променевої терапії. Такі алгоритми дозволяють знаходити комбінації доз і траєкторій опромінення, що забезпечують максимальне знищення пухлинних клітин при мінімальному ушкодженні здорових тканин. Перспективи розвитку генетичних алгоритмів у медицині пов'язані з інтеграцією їх у гібридні системи разом з методами глибокого навчання та іншими еволюційними підходами, що дасть змогу підвищити адаптивність і продуктивність моделей. Одним із викликів залишається висока обчислювальна складність алгоритмів при роботі з дуже великими наборами даних, однак розвиток обчислювальних технологій поступово зменшує цей бар'єр і відкриває нові можливості для широкого застосування генетичних алгоритмів у медицині.

### 2.3.6 Медичні системи на основі NLP

Медичні системи на основі технологій обробки природної мови (Natural Language Processing, NLP) набувають дедалі більшого значення в сучасній медицині завдяки своїй здатності працювати з великими обсягами неструктурованих текстових даних. У клінічній практиці значна частина медичної інформації представлена у вигляді вільного тексту: історії хвороби, протоколи обстежень, консультаційні висновки, рекомендації, епікризи, а також повідомлення від пацієнтів. Технології NLP дають змогу автоматизувати аналіз і структурування таких даних, що підвищує

ефективність роботи лікарів, покращує якість медичної документації та відкриває нові можливості для досліджень і клінічної аналітики.

Основними завданнями NLP у медицині є вилучення сутностей (Named Entity Recognition, NER), класифікація текстів, нормалізація термінів, визначення взаємозв'язків між сутностями та побудова клінічних онтологій. Системи на основі NLP здатні автоматично визначати та маркувати клінічні ознаки, діагнози, лабораторні показники, ліки та дії лікаря у текстах медичних записів. Це дає змогу перетворювати неструктуровану інформацію на структуровані дані, придатні для подальшої обробки й аналізу. Наприклад, завдяки використанню NLP можна автоматично заповнювати медичні бази даних на підставі текстових записів або формувати коди діагнозів за МКХ.

У клінічній практиці NLP використовується для створення систем підтримки прийняття рішень, де алгоритми аналізують текстові записи для виявлення ознак небезпечних станів чи неповних обстежень. Такі системи допомагають лікарям швидше отримати огляд клінічної картини пацієнта й зосередити увагу на критично важливих моментах. Зокрема, у великих медичних центрах США й Європи NLP інтегровано у платформи електронних медичних записів для автоматичного виявлення факторів ризику, генерації нагадувань про необхідність призначення обстежень або вакцинацій. У дослідженнях NLP застосовується для аналізу наукової літератури й клінічних протоколів з метою пошуку актуальної інформації й формування клінічних рекомендацій.

Прикладами успішного застосування NLP у медицині є такі проекти, як cTAKES (clinical Text Analysis and Knowledge Extraction System) – система для обробки клінічних текстів, що активно використовується у дослідницьких і клінічних установах США, та MedLEE (Medical Language Extraction and Encoding), яка була однією з перших систем для автоматичного кодування медичних даних. Іншим прикладом є використання NLP для автоматизації обробки запитів пацієнтів у чат-ботах

і телемедичних платформах, де алгоритми здатні розпізнавати симптоми й формувати попередні рекомендації.

Перспективи розвитку медичних систем на основі NLP пов'язані з поєднанням цих технологій з великими мовними моделями (LLM), що дозволить суттєво підвищити точність і гнучкість обробки клінічних текстів. Крім того, очікується ширше впровадження мультимодальних систем, які комбінуватимуть текстову інформацію з медичними зображеннями, сигналами й іншими даними для формування цілісної клінічної картини. Водночас важливими викликами залишаються питання захисту конфіденційної інформації, стандартизації термінології та інтеграції NLP-рішень у клінічні робочі процеси з урахуванням нормативно-правових вимог. Таким чином, медичні системи на основі NLP відіграють ключову роль у розвитку цифрової медицини та створенні інтелектуальних інструментів для підтримки лікарів і пацієнтів.

### 2.3.7 Медичні системи на основі LLM

Медичні системи на основі великих мовних моделей (Large Language Models, LLM) є одним із найновіших і найдинамічніших напрямів розвитку штучного інтелекту в охороні здоров'я. LLM, такі як GPT (Generative Pre-trained Transformer), BERT (Bidirectional Encoder Representations from Transformers) та їхні медичні модифікації (наприклад, BioGPT, PubMedBERT, Med-PaLM), демонструють здатність аналізувати й генерувати текстові дані з глибиною й точністю, що раніше були недосяжними для автоматизованих систем. У медичному контексті LLM відкривають нові можливості для обробки великих обсягів неструктурованої інформації, підтримки прийняття клінічних рішень, автоматизації документації й навіть взаємодії з пацієнтами через чат-боти й голосових помічників.

Одним із основних завдань, де LLM уже довели свою ефективність, є обробка електронних медичних записів (EMR/EHR). Більша частина клінічної інформації представлена у вигляді вільного тексту – лікарських записів, епікризів, протоколів обстежень, результатів консультацій. LLM здатні автоматично витягати ключові клінічні ознаки, діагнози, результати обстежень, призначення та інші важливі дані. Це дає змогу структурувати інформацію для подальшої аналітики, знижувати адміністративне навантаження на медичних працівників та підвищувати якість ведення медичної документації. Наприклад, моделі, треновані на базі PubMed та клінічних корпусів, демонструють високу точність в автоматичному кодуванні діагнозів за міжнародною класифікацією хвороб (МКХ), формуванні виписок і підготовці звітів.

Ще одним перспективним напрямом використання LLM є створення систем підтримки клінічних рішень. Великі мовні моделі можуть аналізувати симптоми, результати аналізів і попередню історію хвороби пацієнта та пропонувати можливі діагнози або тактики лікування на основі медичних знань. У 2023 році Google представив модель Med-PaLM, натреновану спеціально на медичних даних, яка продемонструвала продуктивність, близьку до експертного рівня в тестах на клінічне мислення. Такі системи можуть використовуватися для допомоги лікарям у формулюванні диференційного діагнозу, перевірки повноти обстеження або навіть для генерування текстів консиліумних висновків.

Окрему нішу займають системи на основі LLM для взаємодії з пацієнтами. Чат-боти й віртуальні помічники з медичним «підґрунтям» можуть відповідати на запитання пацієнтів, надавати попередні поради щодо невідкладних станів, роз'яснювати рекомендації лікаря. Такі рішення вже застосовуються в телемедичних платформах, програмах для моніторингу хронічних захворювань та у сфері ментального здоров'я. Прикладами є Babylon Health (Велика Британія), Ada Health (Німеччина), а також експериментальні системи на базі ChatGPT та його аналогів.

Перспективи розвитку медичних систем на основі LLM пов'язані з інтеграцією їх у багатомодальні системи, що об'єднують текстові дані з результатами зображень, лабораторних тестів і сигналів для більш точного й комплексного аналізу стану пацієнта. Очікується, що такі системи стануть частиною інтелектуальних платформ для персоналізованої медицини та клінічних досліджень. Водночас важливими викликами залишаються питання етики, безпеки й відповідальності, оскільки LLM можуть генерувати помилкові або неоднозначні висновки. Тому актуальним напрямом є розробка механізмів пояснюваності, валідації клінічних рекомендацій та суворе регуляторне врегулювання використання таких технологій у медицині.

### 2.3.8 Нейроінтерфейс «мозок-машина»

Нейроінтерфейси типу «мозок-машина» (Brain-Computer Interface, BCI) є одним із найінноваційніших напрямів розвитку медичних технологій на стику нейронаук, біоінженерії та штучного інтелекту. Їхня основна мета полягає у створенні каналів зв'язку між нервовою системою людини та зовнішніми пристроями без використання м'язів чи периферичної нервової активності. Такі системи здатні безпосередньо зчитувати електричну активність мозку й трансформувати її у керуючі сигнали для протезів, комп'ютерів, роботизованих систем або медичних пристроїв. Нейроінтерфейси відкривають нові можливості для відновлення рухових і когнітивних функцій у пацієнтів із тяжкими неврологічними ураженнями – наприклад, після інсультів, черепно-мозкових травм чи при нейродегенеративних захворюваннях.

Принцип роботи нейроінтерфейсів базується на реєстрації та обробці біоелектричних сигналів мозку. Найчастіше для цього використовуються електроенцефалографія (ЕЕГ), магнітоенцефалографія (МЕГ) або інвазивні методи реєстрації – наприклад, за допомогою імплантованих електродів, які

контактують безпосередньо з корою головного мозку. Ці сигнали проходять попередню обробку для виділення релевантних характеристик – частотних компонентів, просторових шаблонів активності чи специфічних патернів, пов'язаних із виконанням певних ментальних завдань або наміром руху. Далі ці ознаки передаються алгоритмам машинного навчання чи глибоких нейронних мереж, які здійснюють класифікацію або регресію сигналів для генерації керуючих команд.

Одним із найбільш відомих сучасних проєктів у сфері нейроінтерфейсів є Neuralink, створений Ілоном Маском у 2016 році. Метою компанії є розробка високоточних нейроінтерфейсів з імплантованими мікрочипами, які забезпечують двосторонній обмін інформацією між мозком і комп'ютером. Технологія Neuralink передбачає використання гнучких електродних ниток діаметром менше за людську волосину, які імплантуються в кору мозку за допомогою роботизованих хірургічних систем з мінімальною інвазивністю. Основні завдання, які декларує Neuralink, включають відновлення сенсомоторних функцій у паралізованих пацієнтів, можливість керування комп'ютерними інтерфейсами силою думки, а в майбутньому – потенційне лікування хвороби Альцгеймера, Паркінсона, депресії та інших станів шляхом нейромодуляції.

Сучасні прототипи пристроїв Neuralink вже були випробувані на тваринах, де демонструвалася можливість контролювати рухи курсора або простих об'єктів на екрані. У 2024 році Neuralink отримала дозвіл на перші клінічні випробування імплантів на людях. Система інтегрує передові алгоритми обробки сигналів на основі глибокого навчання для реального часу декодування нейрональної активності та генерації керуючих сигналів. У перспективі передбачається створення повноцінних нейроінтерфейсів із високою роздільною здатністю та двостороннім зв'язком, де пристрій зможе не лише зчитувати активність мозку, а й впливати на нього для корекції патологічних станів.

Перспективи розвитку нейроінтерфейсів у медицині пов'язані з удосконаленням бездротових технологій передачі даних, мінімізацією розмірів імплантів і підвищенням їхньої біосумісності. Крім того, великі очікування покладаються на інтеграцію нейроінтерфейсів з іншими технологіями штучного інтелекту для створення адаптивних систем, здатних до самонавчання й персоналізації під потреби конкретного пацієнта. Водночас важливими викликами залишаються етичні та правові аспекти – питання конфіденційності даних, безпеки імплантів, відповідальності за прийняті рішення та впливу на когнітивні функції користувача. Незважаючи на ці виклики, нейроінтерфейси типу «мозок- машина» вважаються однією з найперспективніших технологій для нейрореабілітації, компенсації втрати функцій і, потенційно, розширення когнітивних можливостей людини.

### 2.3.9 Імітаційне моделювання лікарських препаратів

Імітаційне моделювання лікарських препаратів є важливим інструментом сучасної медицини та фармацевтики, що дає змогу досліджувати поведінку лікарських засобів в організмі, оптимізувати їх склад і спосіб застосування, а також прогнозувати ефективність і безпечність терапії ще до початку клінічних випробувань. Цей підхід ґрунтується на створенні математичних і комп'ютерних моделей, які відтворюють фармакокінетику (рух препарату в організмі) і фармакодинаміку (біологічну дію препарату) з урахуванням фізіологічних, біохімічних та генетичних особливостей пацієнтів. Моделювання дозволяє проводити віртуальні експерименти, що значно знижує потребу у тваринних дослідженнях і підвищує ефективність розробки нових лікарських засобів.

Сутність імітаційного моделювання полягає у побудові комп'ютерної репрезентації системи «лікарський засіб – організм», де ключовими компонентами є моделі абсорбції, розподілу, метаболізму та виведення

препарату (ADME-моделі). Для цього використовуються як традиційні статистичні методи, так і сучасні підходи машинного навчання та штучного інтелекту. Значного поширення набули так звані фізіологічно обґрунтовані фармакокінетичні моделі (Physiologically-Based Pharmacokinetic models, PBPK), які дозволяють враховувати анатомо-фізіологічні параметри (об'єм органів, кровоплин, ферментативну активність) та індивідуальні характеристики пацієнтів. Такі моделі дають можливість оцінювати вплив дозування, віку, статі, генетичних варіацій і супутніх захворювань на фармакокінетичний профіль препарату.

Імітаційне моделювання активно використовується на різних етапах життєвого циклу лікарських засобів. На стадії доклінічних досліджень воно дозволяє відбирати найбільш перспективні молекули для подальших випробувань, передбачати можливі токсичні ефекти, підбирати оптимальні шляхи введення та дозування. У процесі клінічних досліджень моделі застосовуються для планування експериментів, створення синтетичних контрольних груп та аналізу результатів випробувань. Крім того, імітаційне моделювання використовується для індивідуалізації фармакотерапії, зокрема у випадках поліпрагмазії або при лікуванні пацієнтів із порушеннями функції печінки чи нирок.

Реальні приклади застосування імітаційного моделювання лікарських засобів можна побачити в роботі таких платформ, як GastroPlus, Simcyp Simulator, PK-Sim, які широко використовуються фармацевтичними компаніями та регуляторними органами. Наприклад, моделювання за допомогою PBPK-систем дозволило оптимізувати схеми дозування ряду онкологічних препаратів і препаратів для лікування ВІЛ, враховуючи вплив лікарських взаємодій. Крім того, у фармакогеноміці імітаційне моделювання допомагає підбирати комбінації препаратів і прогнозувати їхню дію з урахуванням поліморфізмів генів метаболізуючих ферментів.

Перспективи розвитку імітаційного моделювання в медицині та фармацевтиці пов'язані з подальшою інтеграцією цих технологій з

великими мовними моделями та системами штучного інтелекту для автоматичного аналізу результатів моделювання й формування рекомендацій щодо лікування. Очікується також розширення застосування в області персоналізованої медицини, коли моделі будуть створюватися на основі індивідуальних даних конкретного пацієнта для підбору найбільш ефективної й безпечної терапії.

### 3 ОПИС ПРОГРАМНОЇ РЕАЛІЗАЦІЇ

Для створення сучасної інформаційної системи, яка дозволяє автоматично аналізувати медичні дані користувача з використанням методів штучного інтелекту, необхідно реалізувати повноцінний програмний комплекс. Така система складається з декількох основних компонентів, кожен з яких виконує окрему роль у забезпеченні цілісної роботи продукту (рисунок 3.1).

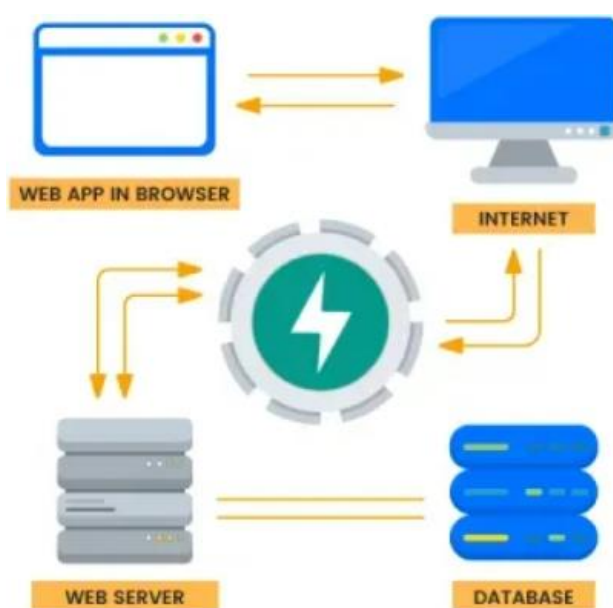


Рисунок 3.1 – Принцип роботи клієнт-серверної архітектури

Frontend також відомий як інтерфейс користувача. Відповідає за взаємодію з користувачем, введення та відображення даних, надання результатів аналізу у зручному вигляді.

Backend або серверна частина. Здійснює обробку запитів від користувача, управління базою даних, виконання алгоритмів машинного навчання та надання результатів аналізу.

Модулі штучного інтелекту. Реалізують моделі для класифікації та прогнозування стану здоров'я на основі аналізів та медичних зображень.

База даних. Забезпечує збереження усіх користувацьких даних, результатів аналізів та історії обстежень.

Зовнішні ресурси та датасети. Використовуються для навчання та тестування моделей штучного інтелекту, а також для поповнення системи довідковою інформацією.

Розглянемо детальніше реалізацію кожного з компонентів.

## 3.1 Frontend

### 3.1.1 HTML

HTML – це основна мова розмітки, яка використовується для створення структури та змісту веб-сторінок. Саме за допомогою HTML визначаються всі елементи інтерфейсу користувача: заголовки, абзаци, таблиці, форми, кнопки, поля вводу тощо. У рамках frontend-розробки HTML формує «скелет» сторінки, на який накладаються стилі та функціональність.

У моєму додатку HTML визначає основну структуру всіх сторінок користувацького інтерфейсу, включаючи дашборд, форму введення аналізів, історію та профіль користувача.

### 3.1.2 CSS

CSS – це мова стилів, яка відповідає за зовнішній вигляд і оформлення веб-сторінок. За допомогою CSS задаються кольори, шрифти, розміри елементів, розташування на сторінці, а також адаптивність інтерфейсу під різні пристрої. CSS дозволяє робити інтерфейс сучасним, зрозумілим та привабливим для користувача.

У моєму проєкті CSS застосовується для оформлення дашборду, кнопок, форм, таблиць результатів аналізів та інших елементів, забезпечуючи зручність та естетичну привабливість додатку.

### 3.1.3 JavaScript

JavaScript – це мова програмування, яка використовується для додавання динамічної поведінки та інтерактивності веб-сторінок. За допомогою JavaScript реалізується обробка подій, валідація форм, анімація, динамічне оновлення змісту сторінки без її повного перезавантаження (технологія AJAX), а також обмін даними з сервером через API.

У моєму додатку JavaScript забезпечує взаємодію користувача з інтерфейсом – наприклад, перевірку правильності введення даних, відправку запитів до backend-сервера для аналізу медичних даних та оновлення результатів аналізів у режимі реального часу.

### 3.2 Backend

Серверна частина системи реалізована на базі фреймворку FastAPI, який поєднує простоту створення REST API з високою продуктивністю завдяки асинхронній обробці запитів.

Backend виконує роль центрального елемента системи, який приймає запити від користувачів, обробляє вхідні дані, викликає відповідну модель машинного навчання чи глибокого навчання та повертає результат клієнту.

Архітектурно серверна частина побудована за принципом клієнт-серверної моделі з чітким поділом логіки обробки даних та модульністю коду.

Кожен сервіс реалізовано у вигляді окремих Python-модулів. Це спрощує масштабування системи та дає можливість легко додавати нові моделі або оновлювати існуючі.

Сервер підтримує роботу з такими форматами вхідних даних, як JSON (для табличних даних) та файли зображень (JPEG, PNG) для рентген-знімків.

### 3.2.1 API

У системі використано FastAPI – сучасний Python-фреймворк для створення веб-інтерфейсів та RESTful API. Його основними перевагами є висока продуктивність завдяки використанню асинхронної моделі обробки запитів, автоматична генерація документації API у форматах OpenAPI та Swagger та зручна інтеграція із засобами валідації даних через Pydantic.

Нижче наведено приклад функції обробки маршруту для аналізу рентген-знімку (рисунок 3.2).

```
@app.post("/predict/pneumonia")
async def predict_pneumonia_api(file: UploadFile = File(...)):
    img_bytes = await file.read()
    result = predict_pneumonia(img_bytes)
    return result
```

Рисунок 3.2 – Функція для обробки запиту

### 3.2.2 Навчання моделей

Навчання моделей у проєкті реалізовано за допомогою популярних Python-бібліотек для обробки даних та машинного навчання [14]. Усі моделі були побудовані, збережені й потім використані у веб-застосунку через API.

Scikit-learn (sklearn). Одна з найпопулярніших бібліотек для класичного машинного навчання. Вона забезпечує великий набір алгоритмів, засоби для оцінки якості моделей та підготовки даних [15].

PyTorch. Бібліотека яка відповідає за створення й навчання глибоких нейронних мереж (у нашому випадку CNN), гнучку роботу з тензорами та можливість використання готових архітектур моделей [16].

У роботі застосовувалась архітектура ResNet18, одна з класичних згорткових нейронних мереж, яка добре зарекомендувала себе при роботі з медичними зображеннями [17]. Модель була попередньо натренована на базі ImageNet. За допомогою модулю torchvision було завантажено готову архітектуру ResNet18 та трансформовано зображення для передобробки перед подачею у модель (рисунок 3.3). Ці трансформації адаптують вхідні дані під формат, на якому була навчена базова модель ResNet18.

```
from torch import nn, optim
from torchvision import transforms, datasets, models

transform = transforms.Compose([
    transforms.Resize((224, 224)),
    transforms.ToTensor(),
    transforms.Normalize([0.485, 0.456, 0.406],
                          [0.229, 0.224, 0.225])
])
model = models.resnet18(pretrained=True)
model.fc = nn.Linear(model.fc.in_features, 2)
```

Рисунок 3.3 – Модуль трансформації зображення та завантаження моделі ResNet18

Після навчання моделі зберігалися у форматі .pkl та .pt за допомогою бібліотеки joblib, спеціалізованої бібліотека для серіалізації та десеріалізації об'єктів Python, яка оптимізована для збереження великих моделей та масивів даних. Вона є стандартом для роботи з моделями у світі машинного навчання, особливо при використанні scikit-learn.

### 3.3 Робота з датасетами

Одним із ключових факторів успішності системи машинного навчання є правильний вибір датасету.

Якість даних безпосередньо впливає на точність, надійність та практичну застосовність побудованих моделей.

У сфері медичних досліджень вибір датасету стає особливо складним завданням тому що є обмежений доступ до медичних даних через їх конфіденційність і захист персональної інформації, висока вартість отримання та обробки даних, що відповідають стандартам доказової медицини, невідповідність багатьох відкритих датасетів вимогам якості.

Під час реалізації даної роботи стояло завдання не просто знайти відкриті датасети, а й обрати ті, що є якісними та підходять для навчання моделей без ризику перенавчання або втрати релевантності результатів.

Diabetes Classification Dataset. Датасет містить дані клінічних показників (вік, ІМТ, рівні холестерину, тригліцеридів та інші параметри), на основі яких модель визначає наявність або відсутність діабету. Дані структуровані та підготовлені для задач класифікації [18].

Heart Failure Prediction Dataset. Цей датасет використано для навчання моделі оцінки ризику серцевої недостатності. Містить ключові клінічні ознаки: вік, фракцію викиду, рівень креатиніну, тромбоцитів тощо. Його перевагою є чистота даних та збалансованість класів [19].

X-ray images (Pneumonia) dataset. Датасет з рентген-знімками грудної клітки для класифікації пневмонії. Містить зображення, поділені на класи: нормальні та із наявністю пневмонії. Його використано для навчання глибокої згорткової нейронної мережі (ResNet18) [20].

### 3.4 База даних

У будь-якому вебзастосунку база даних виконує ключову роль – вона є сховищем для зберігання даних, необхідних для коректної роботи системи.

База даних забезпечує централізоване зберігання інформації про користувачів, накопичення історії запитів і результатів передбачення, а також збереження додаткових налаштувань, журналів активності та журналів помилок.

Для реалізації функцій зберігання та пошуку даних у цій роботі використовується SQLite, систему управління базами даних (СУБД). SQLite відрізняється простотою налаштування та високою продуктивністю.

Схема бази даних включає чотири таблиці (рисунок 3.4):

- таблиця `users`, містить реєстраційну інформацію користувачів;
- таблиця `physical_params`, зберігає інформацію про фізичні параметри пацієнта;
- таблиця `lab_test`, зберігає файли аналізів які завантажують користувачі;
- таблиця `examination_results`, результати діагностики захворювань системою.

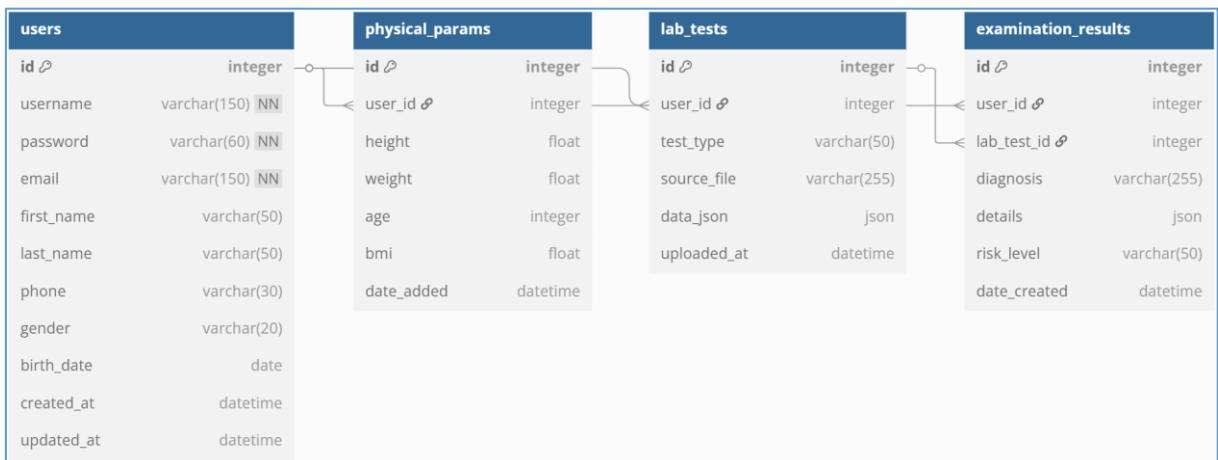


Рисунок 3.4 – Схема бази даних

Опис таблиці users:

– id, унікальний ідентифікатор користувача. Має тип даних integer. Обмеження первинного ключа гарантує, що кожен користувач буде мати унікальний ідентифікатор;

– username, логін користувача. Має тип даних varchar. Обмеження unique гарантує, що імена користувачів відрізняються, запобігаючи створенню дублікатів. Обмеження not null гарантує, що це поле обов'язково має бути заповнено;

– password, хеш пароля користувача. Має тип даних varchar. Обмеження not null гарантує, що це поле обов'язково має бути заповнено;

– email, адреса електронної пошти. Має тип даних varchar. Обмеження unique гарантує, що імена користувачів відрізняються, запобігаючи створенню дублікатів. Обмеження not null гарантує, що це поле обов'язково має бути заповнено;

– first\_name, ім'я користувача. Має тип даних varchar;

– last\_name, прізвище користувача. Має тип даних varchar;

– phone, номер телефону. Має тип даних varchar;

– gender, стать користувача. Має тип даних varchar;

– birth\_date, дата народження користувача. Має тип даних date;

– created\_at, дата створення облікового запису. Має тип даних datetime;

– updated\_at, дата останнього оновлення запису. Має тип даних datetime.

Опис таблиці physical\_params:

– id, унікальний ідентифікатор запису. Має тип даних integer. Обмеження первинного ключа гарантує унікальність кожного запису фізичних параметрів;

– user\_id, ідентифікатор користувача, якому належать ці фізичні параметри. Має тип даних integer. Є зовнішнім ключем, що посилається на

поле `id` таблиці `users`, забезпечуючи зв'язок між користувачем і його параметрами;

- `height`, зріст користувача. Має тип даних `float`;
- `weight`, вага користувача. Має тип даних `float`;
- `age`, вік користувача. Має тип даних `integer`;
- `bmi`, індекс маси тіла користувача. Має тип даних `float`;
- `date_added`, дата додавання запису. Має тип даних `datetime`, що

відображає час внесення параметрів до бази.

Опис таблиці `lab_tests`:

- `id`, унікальний ідентифікатор лабораторного тесту. Має тип даних `integer`. Первинний ключ забезпечує унікальність запису;

- `user_id`, ідентифікатор користувача, який завантажив або якому належить цей тест. Має тип даних `integer`. Є зовнішнім ключем, що посиляється на `users.id`;

- `test_type`, тип лабораторного аналізу. Має тип даних `varchar`;
- `source_file`, назва або шлях до завантаженого файлу з даними тесту.

Має тип даних `varchar`;

- `data_json`, дані тесту у форматі JSON, що зберігають результати у структурованому вигляді. Має тип даних `json`;

- `uploaded_at`, дата завантаження тесту. Має тип даних `datetime`, вказує час додавання даних у систему.

Опис таблиці `examination_results`:

- `id`, унікальний ідентифікатор результату обстеження. Має тип даних `integer`. Первинний ключ забезпечує унікальність запису;

- `user_id`, ідентифікатор користувача, якому належить результат. Має тип даних `integer`. Є зовнішнім ключем, що посиляється на `users.id`;

- `lab_test_id`, ідентифікатор лабораторного тесту, що використовувався для формування результату. Має тип даних `integer`, зовнішній ключ на `lab_tests.id`;

- `diagnosis`, основний діагноз, сформований на основі результатів обстеження. Має тип даних `varchar`;
- `details`, додаткові дані у форматі JSON (наприклад, ймовірності для кількох діагнозів). Має тип даних `json`;
- `risk_level`, рівень ризику, визначений системою. Має тип даних `varchar`;
- `date_created`, дата створення запису результату. Має тип даних `datetime`, фіксує момент генерації результату.

### 3.5 Логіка роботи системи

Логіка роботи системи побудована за принципом клієнт-серверної архітектури, де `frontend` виступає як клієнтський інтерфейс, а `backend` (`FastAPI`) – як серверна частина, що обробляє запити та взаємодіє з моделями й базою даних.

Користувач заходить у вебінтерфейс системи, авторизується та завантажує дані клінічних аналізів.

`Frontend` формує запит (у форматі JSON) і надсилає його до бекенду по HTTPS.

`FastAPI` за допомогою моделей `Pydantic` автоматично перевіряє структуру запиту, обирає потрібний маршрут та викликає відповідний сервіс для підготовки даних і виклику моделі.

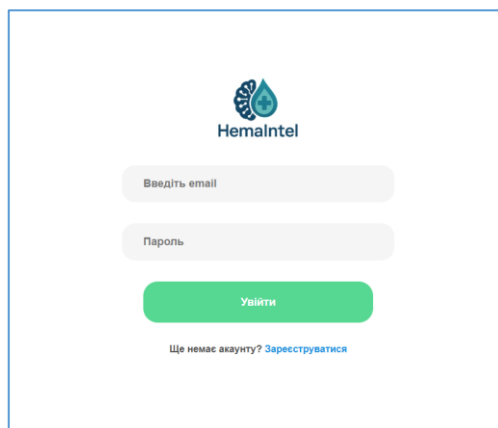
Сервіс виконує передобробку (масштабування, перетворення, трансформацію тощо) вхідних даних і викликає модель для аналізу цих даних. Модель обчислює прогноз і повертає результат (класифікацію, ймовірності).

Результати прогнозу записуються у базу даних. створюється запис у таблиці `examination_results`, зберігаються додаткові дані, фіксується зв'язок між результатом, користувачем і лабораторним тестом.

`FastAPI` формує відповідь у форматі JSON і повертає її клієнту.

## 4 ІНТЕРФЕЙС КОРИСТУВАЧА

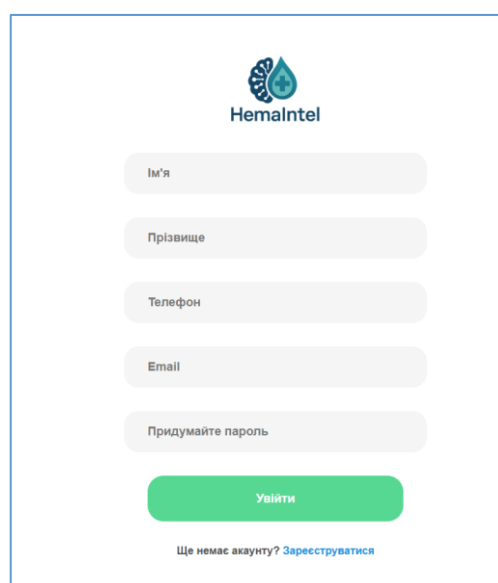
Користувач переходить на сайт і потрапляє на сторінку авторизації (рисунок 4.1).



The screenshot shows the login interface for Hemaintel. At the top center is the Hemaintel logo, which consists of a blue water drop icon with a leaf inside, and the text 'Hemaintel' below it. Below the logo are two input fields: 'Введіть email' and 'Пароль'. Below these fields is a green button labeled 'Увійти'. At the bottom of the form, there is a link that says 'Ще немає акаунту? Зареєструватися'.

Рисунок 4.1 – Сторінка авторизації

Йому пропонується ввести дані для входу. Якщо у користувача немає зареєстрованого акаунту, йому пропонується перейти на сторінку реєстрації (рисунок 4.2).



The screenshot shows the registration interface for Hemaintel. At the top center is the Hemaintel logo, which consists of a blue water drop icon with a leaf inside, and the text 'Hemaintel' below it. Below the logo are five input fields: 'Ім'я', 'Прізвище', 'Телефон', 'Email', and 'Придумайте пароль'. Below these fields is a green button labeled 'Увійти'. At the bottom of the form, there is a link that says 'Ще немає акаунту? Зареєструватися'.

Рисунок 4.2 – Сторінка реєстрації

При спробі входу система перевіряє правильність логіна та пароля у базі даних. Якщо дані введено некоректно, користувач бачить повідомлення про помилку та може повторити спробу. У разі успішної авторизації користувач потрапляє на головну сторінку додатку (рисунок 4.3).

**Огляд**  
Не ігноруйте проблеми зі здоров'ям. Лікування захворювань на ранніх стадіях запобігає ускладненням.

**Статус здоров'я**

Згідно з останніми аналізами у вас підвищений шанс інфаркту

На рентгені легень патологій не виявлено

**Результати останнього обслідування** [Детальніше](#)

|  |     |         |         |
|--|-----|---------|---------|
| Загальний холестерин                                     | 6,1 | 5,2     | ммоль/л |
| Ліпопротеїни високої щільності(ЛПВЩ)                     | 0,9 | >1      | ммоль/л |
| Ліпопротеїни низької щільності (бета-ліпопротеїни, ЛПНЩ) | 4,2 | 2,5-3,3 | ммоль/л |
| Тригліцериди   | 2,2 | до 2,3  | ммоль/л |

**Останні аналізи** [Переглянути історію](#)

- Зозуля\_100768804\_Біохімічний\_аналіз\_крові  
27.04.2025
- Зозуля\_100768402\_Загальний\_аналіз\_крові  
12.03.2025
- Зозуля\_100756262\_Рентген\_легень  
14.11.2024

Рисунок 4.3 – Головна сторінка

На головній сторінці він бачить навігаційне меню та загальну інформацію про свій обліковий запис. Якщо потрібно додати новий аналіз, користувач переходить на сторінку «Додати аналізи» (рисунок 4.4), де обирає тип даних, заповнює форму або завантажує файл і відправляє дані на обробку.

Після обробки на сторінці автоматично відображається результат, де користувач може побачити основний діагноз, додаткові дані та рівень ризику (рисунок 4.5).

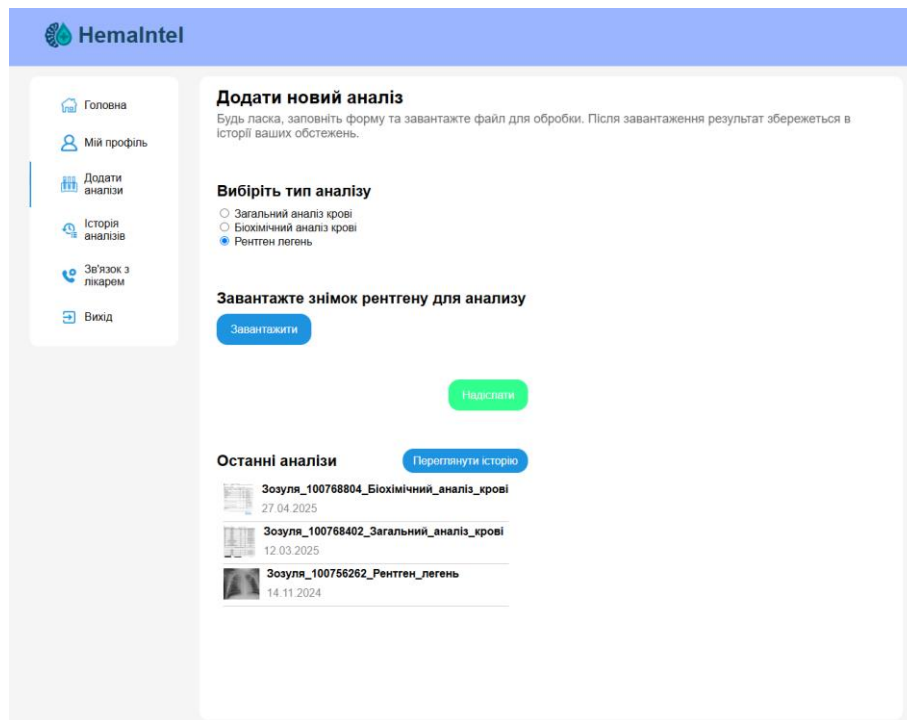


Рисунок 4.4 – Сторінка «Додати аналізи»

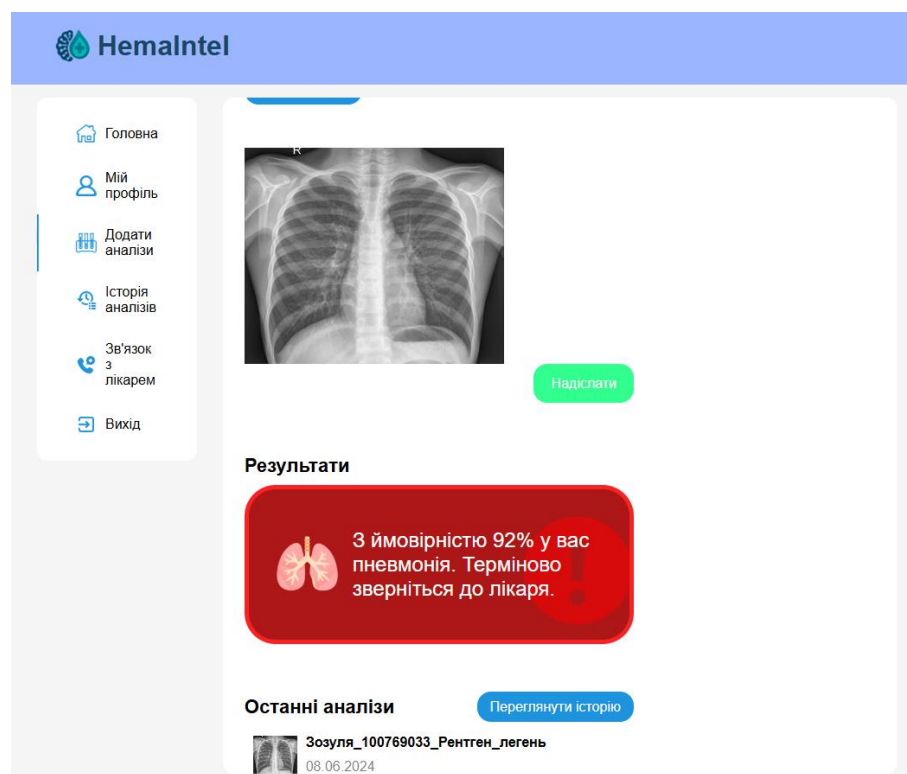


Рисунок 4.5 – Результат аналізу

## ВИСНОВКИ

У результаті виконання кваліфікаційної роботи було розроблено вебзастосунок для діагностики захворювань засобами штучного інтелекту, що дозволяє здійснювати попередню діагностику на основі табличних даних та рентген-зображень. Завдання кваліфікаційної роботи виконано повністю. Реалізовано архітектуру системи, створено модулі для прийому й обробки даних, інтегровано моделі машинного навчання та глибокого навчання для прогнозування ризику діабету, серцевої недостатності та діагностики пневмонії. Досягнуті якісні показники моделей відповідають сучасним вимогам до подібних систем: точність моделей машинного навчання перевищує 80 %, CNN-модель для діагностики пневмонії демонструє високу якість класифікації з точністю понад 90 % на тестових даних.

Розроблена система є конкурентною порівняно з вітчизняними й світовими аналогами завдяки гнучкій архітектурі, можливості легко додавати нові моделі та типи даних, а також використанню відкритих стандартів і бібліотек (FastAPI, scikit-learn, PyTorch). У світовій практиці подібні рішення реалізовані в продуктах таких компаній, як IBM (Watson for Health), Google (DeepMind Health). Водночас запропонована система відрізняється відкритістю, можливістю налаштування під конкретні потреби користувача й оптимізацією під обмежені обчислювальні ресурси.

Варто відзначити, що розробка таких систем супроводжується значною проблематикою, зокрема складністю пошуку відкритих і якісних анонімізованих медичних датасетів. Світовому науковому та медичному співтовариству необхідно сприяти поширенню й розвитку відкритих анонімізованих баз даних, що дозволить прискорити впровадження систем штучного інтелекту в медичну практику. Окрім цього, при інтеграції подібних систем важливо враховувати етичні аспекти, зокрема збереження конфіденційності даних пацієнтів, уникнення дискримінації на основі

результатів прогнозування та обов'язкову роль лікаря у прийнятті остаточного клінічного рішення.

Розробка тісно пов'язана з науково-дослідними розробками кафедри штучного інтелекту Харківського національного університету радіоелектроніки, зокрема з напрямками створення інтелектуальних систем обробки медичних даних, що знайшло відображення у публікаціях кафедри. Також результати роботи базуються на дослідженнях, які були представлені в науковій роботі автора на тему «Діагностика захворювань по аналізам крові засобами ШІ» (2025 р.), де розглядалися питання стандартизації та уніфікації медичних даних, застосування глибинного навчання для обробки лабораторних показників [21].

У ході роботи отримано нові наукові результати щодо побудови архітектури вебзастосунку для медичної діагностики, організації бази даних пацієнтів та інтеграції моделей ШІ з інтерфейсом користувача. Ці результати можуть бути використані для подальшого розширення системи, зокрема для додавання моделей для прогнозування інших захворювань, розробки модуля лікарського кабінету та реалізації персоналізованих рекомендацій.

Матеріали цієї кваліфікаційної роботи можуть бути використані для подальшого розвитку навчального процесу на кафедрі штучного інтелекту Харківського національного університету радіоелектроніки під час підготовки фахівців з напрямку комп'ютерних наук і штучного інтелекту. Зокрема, розроблений застосунок може бути використаний як навчальний приклад при вивченні основ машинного навчання, веброзробки та побудови інтелектуальних систем.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Lundervold A. S., Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*. 2019. Т. 29, № 2. С. 102–127. URL: <https://doi.org/10.1016/j.zemedi.2018.11.002> (дата звернення: 01.06.2025).
2. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* / ed. by H. Trevor, T. Robert, S. (. service). New York, NY : Springer-Verlag New York, 2009. 745 p.
3. Rajpurkar P., Irvin J., Zhu K. et al. CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint. 2017. arXiv:1711.05225.
4. Strodthoff N., Wagner P., Schaeffter T., Samek W. Deep learning for ECG analysis: Benchmarks and insights from PTB-XL. *IEEE J. Biomed. Health Inform.* 2021. Vol. 25(5). P. 1519–1528.
5. ChestX-ray14 dataset. Papers with Code. URL: <https://paperswithcode.com/dataset/chestx-ray14> (дата звернення: 02.06.2025).
6. Goodfellow I., Bengio Y., Courville A. *Deep learning*. Cambridge, MA: MIT Press, 2016. 800 p.
7. Dermatologist-level classification of skin cancer with deep neural networks / A. Esteva та ін. *Nature*. 2017. Т. 542, № 7639. С. 115–118. URL: <https://doi.org/10.1038/nature21056> (дата звернення: 06.06.2025).
8. Rajkomar A., Dean J., Kohane I. Machine learning in medicine. *New England Journal of Medicine*, 2019; 380:1347-1358.
9. Bishop C. M. *Pattern recognition and machine learning*. New York: Springer, 2006. 738 p.
10. Obermeyer Z., Emanuel E. J. Predicting the Future – Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*. 2016. Vol. 375, no. 13. P. 1216–1219. URL: <https://doi.org/10.1056/nejmp1606181> (дата звернення: 09.06.2025).

11. Lipton Z.C., Kale D.C., Elkan C., Wetzel R. Learning to diagnose with LSTM recurrent neural networks. *arXiv preprint*, 2015. <https://arxiv.org/abs/1511.03677> (дата звернення: 12.05.2025)
12. Multitask learning and benchmarking with clinical time series data / Н. Harutyunyan et al. *Scientific Data*. 2019. Vol. 6, no. 1. URL: <https://doi.org/10.1038/s41597-019-0103-9> (дата звернення: 20.05.2025).
13. Miller R. A. Medical Diagnostic Decision Support Systems--Past, Present, And Future: A Threaded Bibliography and Brief Commentary. *Journal of the American Medical Informatics Association*. 1994. Т. 1, № 1. С. 8–27. URL: <https://doi.org/10.1136/jamia.1994.95236141> (дата звернення: 03.06.2025).
14. Chollet F. *Deep learning with Python*. 2nd ed. Shelter Island, NY: Manning Publications, 2021. 504 p.
15. Pedregosa F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 2011. Vol. 12. P. 2825–2830.
16. Paszke A. et al. PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*. 2019. Vol. 32.
17. Shorten C., Khoshgoftaar T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*. 2019. Vol. 6. Article 60.
18. Diabetes Classification Dataset. URL: <https://www.kaggle.com/datasets/simaanjali/diabetes-classification-dataset> (дата звернення: 12.05.2025).
19. Heart Failure Prediction Dataset. URL: <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction> (дата звернення: 14.05.2025).
20. Chollet F. X-ray images (Pneumonia) dataset. URL: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia> (дата звернення: 02.06.2025).

21. Зозуля, Н. О., Гриньова, О. Є. Діагностика захворювань по аналізам крові засобами ШІ. *Радіоелектроніка та молодь у XXI столітті*, м. Харків. 2025. С. 29-30.  
URL: <https://openarchive.nure.ua/handle/document/30693> (дата звернення: 10.06.2025).

