

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

Галузь знань _____ 12 – Інформаційні технології _____

Спеціальність _____ 126 – Інформаційні системи та технології _____

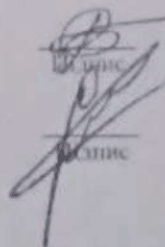
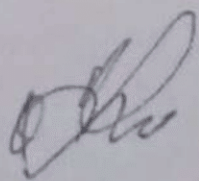
на тему: «Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу»

КвРІСТ. 2301124.23.01.06 ПЗ

Виконав: студент 2 курсу, група ІСТм-23-1

Керівник: к.т.н., доцент
Науковий ступінь, вчене звання

До захисту допускаю:
В.о. зав. кафедри КІС,
PhD Ольга ПАВЛОВА
11 12 2024 р.



Владислав КУЛЬБАЧНИЙ
Ініціал, прізвище

Андрій НІЧЕПОРУК
Ініціал, прізвище

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ

Освітній рівень МАГІСТР

Галузь знань 12 ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

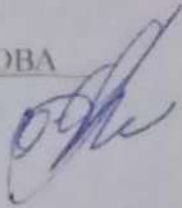
Спеціальність 126 ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

Освітня програма ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА «ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ»

ЗАТВЕРДЖУЮ

В.о. зав. кафедри Ольга ПАВЛОВА

“ 01 ” 11 2024 р.



ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА

Кульбачний Владислав Васильович

Прізвище, ім'я, по батькові студента

1. Тема проекту (роботи) Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Керівник проекту (роботи) Андрій ПІЧЕПОРУК, к.т.н., доцент

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 26.08.2024 р. № 60

2. Строк подання студентом проекту (роботи) на кафедру 01.12.2024 р.

3. Вихідні дані до проекту (роботи) Завдання на дипломне проектування

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)

Аналіз відомих засобів та методів підтримки прийняття рішення для визначення медичного діагнозу

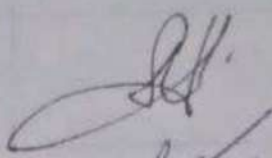
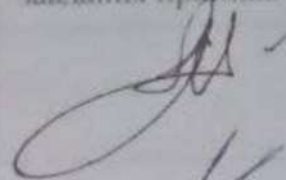


Модель процесу діагностики ішемічної хвороби серця

Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Оцінка ефективності інформаційної системи підтримки прийняття рішення для визначення медичного діагнозу

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень)

6. Консультанти розділів кваліфікаційної роботи магістра

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Нормоконтроль	Сергій ЛИСЕНКО, професор кафедри КІС		
Антиплагіат	Андрій НІЧЕПОРУК, доцент кафедри КІС		

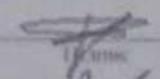
7. Дата видачі завдання « 01 » 09 2024р.

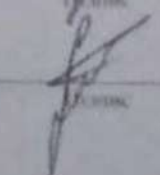
КАЛЕНДАРНИЙ ПЛАН

№з/п	Назва етапів (розділів) дипломного проекту (роботи)	Термін виконання етапів проекту (роботи)	Примітка
1	Вибір напрямку дослідження та узгодження тематики кваліфікаційної роботи з керівником	01.09.2024	виконано
2	Ознайомлення з предметною областю; формулювання мети та задач дослідження; визначення об'єкта та предмета дослідження	15.09.2024	виконано
3	Робота над розділом 1 – дослідження предметної області та постановка задачі	01.10.2024	виконано
4	Робота над розділом 2 – розробка моделей для вирішення поставленої задачі	15.10.2024	виконано
5	Робота над науковою публікацією	15.10.2024	виконано
6	Робота над розділом 3 – розроблення методів для вирішення поставленої задачі	01.11.2024	виконано
7	Робота над розділом 4 – проектування та розробка ПЗ для вирішення поставленої задачі	15.11.2024	
8	Оформлення пояснювальної записки згідно вимог	01.12.2024	виконано
9	Попередній захист ВКР	02.12.2024	виконано
10	Захист ВКР на засіданні ЕК	До 20.12.2024	

Студент

Керівник роботи


Підпис


Підпис

Владислав КУЛЬБАЧНИЙ

Підпис, прізвище

Андрій НІЧЕПОРУК

Підпис, прізвище

РЕФЕРАТ

Тема кваліфікаційної роботи магістра: Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу.

Автор роботи: Владислав КУЛЬБАЧНИЙ

Керівник роботи: Андрій НІЧЕПОРУК, доцент кафедри КІС

Пояснювальна записка: 74 с., 11 рис., 8 табл., 2 дод., 74 джерело.

Перелік ключових слів: медична діагностика, підтримка прийняття рішень, інтеграція систем, аналіз даних.

Об'єктом дослідження є процес діагностики медичних станів з використанням інформаційних систем підтримки прийняття рішень.

Предметом дослідження є методи і моделі машинного навчання, які використовуються для розробки системи підтримки прийняття рішення, спрямованої на точне встановлення медичного діагнозу.

Метою кваліфікаційної є підвищення достовірності формування медичного діагнозу шляхом розробки інформаційної системи підтримки прийняття рішення для діагностики медичних станів.

Методи дослідження. У роботі було застосовано наступні теорії та засоби:

- аналітичні та математичні методи дослідження;
- засоби комп'ютерних мереж;
- методи оцінки ефективності;
- сучасні програмні засоби проектування та дослідження.

Наукова новизна роботи:

Набула подальшого розвитку інформаційна система підтримки прийняття рішення для визначення медичного діагнозу, яка на відміну від відомих залучає адаптивні алгоритми машинного навчання для аналізу великих обсягів медичних даних, що дозволило підвищити достовірність діагностичних процесів

Набула подальшого розвитку модель процесу генерації нечітких правил, яка, на відміну від відомих, включає значення підтримки правил та їх ступінь достовірності, що дозволило враховувати не лише точність класифікації, а й частоту виникнення правил у навчальних даних, що зменшило ймовірність помилкових узагальнень.

Практична цінність роботи полягає у створенні прототипу інформаційної системи, яка підвищує ефективність та достовірність діагностики ішемічної хвороби серця за допомогою сучасних методів машинного навчання. Результати роботи можуть бути використані для впровадження подібних систем у медичну практику.

У першому розділі проведено аналіз сучасних підходів до побудови інформаційних систем підтримки прийняття рішень у медичній діагностиці. Особливу увагу приділено їхньому застосуванню для діагностики серцево-судинних захворювань.

У другому розділі розглянуто методи машинного навчання, які були обрані для діагностики ішемічної хвороби серця. Детально описано процес побудови моделі діагностики, включаючи етапи збору та попередньої обробки даних.

У третьому розділі представлено архітектуру та реалізацію прототипу інформаційної системи. Розроблено програмний модуль, який інтегрує методи машинного навчання для автоматизації процесу діагностики.

У четвертому розділі наведено результати експериментальних досліджень. Оцінено достовірність та загальну ефективність системи на основі тестових даних, а також виконано аналіз її практичної застосовності.

ЗМІСТ

ВСТУП	6
1 АНАЛІЗ ВІДОМИХ ЗАСОБІВ ТА МЕТОДІВ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ	11
1.1 Аналіз предметної області та виявлення наявних проблем і завдань	11
1.2 Аналіз методів для розпізнавання ішемічної хвороби серця	13
1.3 Методи м'яких обчислень у системах діагностики	16
1.4 Опис відомих наукових методів і систем	17
1.5 Опис даних та методів їх обробки.....	20
1.6 Постановка задачі дослідження.....	22
1.7 Висновки	25
2 МОДЕЛЬ ПРОЦЕСУ ДІАГНОСТИКИ ІШЕМІЧНОЇ ХВОРОБИ СЕРЦЯ	27
2.1 Етапи розробки та методологія.....	27
2.2 Архітектура інформаційної системи для діагностики ІХС.....	28
2.3 Модель процесу діагностики ішемічної хвороби серця.....	31
2.4 Моделі машинного навчання для діагностики ішемічної хвороби серця..	33
2.4.1 Математичні основи методів машинного навчання.....	36
2.4.2 Порівняння моделей для діагностики ІХС.....	37
2.4.3 Налаштування параметрів.....	37
2.4.4 Обмеження і способи їх подолання	38
2.4.5 Практичне використання у діагностиці ІХС	38
2.5 Висновки	39
3 ІНФОРМАЦІЙНА СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ	41

3.1 Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу	41
3.2 Метод визначення медичного діагнозу.....	46
3.2.1 Вибір ознак	46
3.2.2 Модель процесу генерації нечітких правил	53
3.2.3 Процедура генерації нечітких правил.....	55
3.2.4 Оптимізація системи на основі нечітких правил.....	58
3.2.5 Оцінка системи на основі нечітких правил	61
3.3 Висновки	62
4 ОЦІНКА ЕФЕКТИВНОСТІ ІНФОРМАЦІЙНОЇ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ	64
4.1 Постановка експериментальних досліджень.....	64
4.2 Процедура вибору ознак.....	65
4.3 Нечітка оптимізація	66
4.3.1 Критерій прозорості.....	67
4.3.2 Критерій достовірності класифікації	68
4.3.3 Чутливість і специфічність	70
4.4 Лінгвістичні нечіткі правила підмножини 2	72
4.5 Інтерфейс інформаційної системи для діагностики	73
4.6 Візуалізація діагностичних даних	76
ВИСНОВКИ	81
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ	84
ДОДАТОК А КОПІЯ ОПУБЛІКОВАНОЇ НАУКОВОЇ СТАТТІ.....	91
ДОДАТОК Б ПРЕЗЕНТАЦІЯ ДО ЗАХИСТУ	97
ДОДАТОК В РЕЗУЛЬТАТИ.....	105

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ІСППР – Інформаційна система підтримки прийняття рішення

СППР – Система підтримки прийняття рішення

НСНП – Нечітка система на основі правил

ІХС – Ішемічна хвороби серця

САК – Стратегія ансамблевих класифікаторів

ШНМ – Штучні нейронні мережі

МОВ – Метод опорних векторів

ВСТУП

Розвиток сучасної медицини невіддільний від швидкого технологічного прогресу, який призводить до появи нових методів обробки інформації та аналізу даних, що мають потенціал кардинально змінити підходи до діагностики та лікування захворювань. Одним із важливих напрямів цієї інтеграції є впровадження інформаційних систем підтримки прийняття рішень (ІСППР) у клінічну практику. Такі системи базуються на передових алгоритмах обробки великих обсягів медичних даних та застосуванні штучного інтелекту, що дозволяє надавати лікарям додаткові інструменти для аналізу та оцінки стану пацієнтів. Основною метою ІСППР є підвищення достовірності діагностики, зменшення ризику помилок та оптимізація часу прийняття клінічних рішень, що має особливе значення у випадках тяжких або хронічних захворювань.

Однією з основних загроз сучасному суспільству є ішемічна хвороба серця (ІХС) — поширене хронічне захворювання, спричинене порушенням кровопостачання серцевого м'яза через звуження коронарних артерій. Це призводить до значного ризику виникнення таких ускладнень, як інфаркт міокарда, серцева недостатність та раптова серцева смерть. ІХС є основною причиною смертності у всьому світі, особливо в країнах із розвиненими економіками, де рівень стресу та малорухливий спосіб життя сприяють поширенню цього захворювання. Отже, своєчасна та точна діагностика ІХС є пріоритетним завданням, що дозволяє знизити ризик тяжких наслідків та покращити якість життя пацієнтів.

Однак процес діагностики ІХС супроводжується значними викликами. Перш за все, клінічні симптоми можуть бути нечіткими або нагадувати ознаки інших захворювань, що ускладнює швидке встановлення правильного діагнозу. Крім того, інвазивні методи, такі як коронарна ангиографія, хоч і забезпечують високу достовірність, є дорогими та мають ризики для пацієнта. Неінвазивні методи, такі як електрокардіографія з фізичним навантаженням, можуть

виявляти зміни на пізніх стадіях або бути менш ефективними на ранніх стадіях хвороби. Це створює потребу у пошуку нових рішень, здатних забезпечити точну діагностику без значних витрат і ризиків для здоров'я пацієнтів.

На сьогодні зростає інтерес до розробки автоматизованих систем, що здатні допомогти лікарям у процесі діагностики ІХС. Завдяки використанню сучасних методів обробки даних, таких як штучні нейронні мережі, машинне навчання та методи аналізу великих даних, стає можливим виявлення прихованих закономірностей у медичних даних, які залишаються непомітними для лікарів у рамках традиційного підходу. Алгоритми машинного навчання дозволяють виявляти взаємозв'язки між клінічними показниками, аналізувати історію хвороби пацієнта та на основі цього прогнозувати ризики розвитку ІХС. Такі можливості допомагають лікарям приймати обґрунтовані рішення про доцільність додаткових діагностичних процедур або своєчасне призначення профілактичних заходів і лікування.

Таким чином, використання ІСППР у діагностиці ІХС є важливим кроком до підвищення ефективності медичної допомоги та зниження смертності від цього захворювання. Розробка таких систем є актуальною і важливою задачею, яка потребує міждисциплінарного підходу та об'єднання зусиль фахівців у галузях медицини, інформатики та інженерії.

Актуальність використання ІСППР у медичній практиці важко переоцінити, особливо в умовах стрімкого зростання обсягів медичних даних і високої захворюваності на ІХС. Ускладнення цієї хвороби вимагають точного та своєчасного діагностування, оскільки затримка в прийнятті рішення може призвести до незворотних наслідків для пацієнта. Застосування інформаційних систем підтримки прийняття рішень може значно покращити ефективність діагностики, зокрема шляхом інтеграції різних джерел медичної інформації та використання алгоритмів для її автоматичної обробки.

ІХС є класичним прикладом захворювання, яке потребує комплексного підходу до діагностики. Різноманітність проявів хвороби, залежно від стадії та особливостей організму пацієнта, ускладнює діагностичний процес.

Використання інформаційних систем для підтримки прийняття рішень дозволяє лікарям проводити більш точний аналіз симптомів, використовуючи великі обсяги даних, включаючи історію хвороби, результати аналізів, електрокардіограми та інші діагностичні дані. Це дозволяє мінімізувати ризики помилкової діагностики та приймати більш точні рішення на основі алгоритмів машинного навчання.

Метою цієї роботи є розробка інформаційної системи підтримки прийняття рішень для діагностики ішемічної хвороби серця, що базується на сучасних алгоритмах обробки даних та машинного навчання. Ця система повинна допомогти підвищити достовірність діагностики, зменшити час, необхідний для прийняття рішень, а також оптимізувати процеси обробки медичних даних.

Для досягнення цієї мети необхідно виконати такі завдання:

1. Провести аналіз існуючих інформаційних систем для підтримки прийняття рішень у діагностиці ІХС.
2. Дослідити сучасні методи обробки медичних даних та алгоритми машинного навчання, що використовуються для діагностики захворювань серцево-судинної системи.
3. Розробити архітектуру інформаційної системи, що здатна ефективно обробляти дані пацієнтів з ІХС.
4. Реалізувати прототип системи у вигляді програмного забезпечення та провести його тестування на реальних медичних даних.
5. Оцінити ефективність запропонованої системи на основі експериментальних досліджень та визначити можливі шляхи її подальшого вдосконалення.

Об'єктом дослідження є процес підтримки прийняття рішень у діагностиці ішемічної хвороби серця. Предметом дослідження є інформаційні системи підтримки прийняття рішень, методи машинного навчання та алгоритми обробки медичних даних, що застосовуються для діагностики ІХС.

Наукова новизна цієї роботи полягає у розробці нового підходу до автоматизації процесів діагностики ішемічної хвороби серця за допомогою сучасних методів обробки даних. Запропонована система підтримки прийняття рішень використовує адаптивні алгоритми машинного навчання для аналізу великих обсягів медичних даних, що дозволяє підвищити достовірність діагностичних процесів. На відміну від існуючих підходів, система, розроблена в рамках цього дослідження, орієнтована на забезпечення високої гнучкості в адаптації до різних типів медичних даних і умов діагностики, що робить її універсальним інструментом для використання в різних медичних закладах.

Практична значимість роботи полягає в можливості впровадження розробленої системи в реальні медичні заклади для автоматизації процесів діагностики ішемічної хвороби серця. Програмне забезпечення, розроблене на основі цієї системи, дозволить лікарям швидше і точніше приймати рішення щодо стану пацієнтів, що значно підвищить якість медичного обслуговування.

Кваліфікаційна робота складається з вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. У першому розділі представлено аналіз сучасних підходів до побудови інформаційних систем підтримки прийняття рішень у медичній діагностиці. Другий розділ присвячений вибору методів машинного навчання для діагностики ІХС. У третьому розділі описано архітектуру та реалізацію прототипу інформаційної системи. У четвертому розділі наведено результати експериментальних досліджень і оцінку ефективності системи. Висновки містять основні результати роботи та пропозиції щодо її подальшого вдосконалення.

Набула подальшого розвитку інформаційна система підтримки прийняття рішення для визначення медичного діагнозу, яка на відміну від відомих залучає адаптивні алгоритми машинного навчання для аналізу великих обсягів медичних даних, що дозволило підвищити достовірність діагностичних процесів

Набула подальшого розвитку модель процесу генерації нечітких правил, яка, на відміну від відомих, включає значення підтримки правил та їх ступінь

достовірності, що дозволило враховувати не лише достовірність класифікації, а й частоту виникнення правил у навчальних даних, що зменшило ймовірність помилкових узагальнень.

1 АНАЛІЗ ВІДОМИХ ЗАСОБІВ ТА МЕТОДІВ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ

1.1 Аналіз предметної області та виявлення наявних проблем і завдань

Ішемічна хвороба серця (ІХС) є провідною причиною смертності в усьому світі, що робить її одним із найбільш важливих напрямів для вдосконалення методів діагностики та профілактики. ІХС виникає внаслідок зниження кровопостачання серцевого м'яза через звуження коронарних артерій, яке спричинене накопиченням бляшок. Це призводить до значних ускладнень, таких як стенокардія, інфаркт міокарда та серцева недостатність. Діагностика ІХС є складною через необхідність оцінки численних факторів, включаючи симптоми, анамнез, результати лабораторних тестів і зображень медичної візуалізації.

Сучасні інформаційні технології та методи обробки великих обсягів даних дають можливість оптимізувати процес діагностики ІХС, покращуючи достовірність діагностичних рішень. Використання автоматизованих інформаційних систем підтримки прийняття рішень (ІСППР) у сфері медичної діагностики дозволяє лікарям ефективніше аналізувати наявні дані та приймати обґрунтовані рішення. У цьому розділі розглянуто ключові проблеми, з якими стикається сучасна медицина у діагностиці ІХС, та завдання, які мають бути вирішені для підвищення ефективності та доступності діагностичних процесів.

Основні проблеми в діагностиці ІХС:

1. Необхідність обробки великих обсягів медичних даних. Медична інформація включає різноманітні дані, такі як результати ЕКГ, УЗД серця, лабораторні показники та симптоми. Обробка таких обсягів інформації вимагає ефективних алгоритмів для систематизації та аналізу.

2. Недостатня достовірність існуючих діагностичних методів. Визначення наявності та ступеня ІХС потребує комплексного підходу, де можливі варіації у симптомах ускладнюють діагностику. Неправильний або

неточний діагноз може призвести до необґрунтованого лікування, що потенційно підвищує ризик серйозних ускладнень.

3. Обмеженість ресурсів для проведення інвазивних досліджень. Золотим стандартом діагностики ІХС є інвазивна коронарна ангіографія, проте цей метод дорогий і технічно складний. Це призводить до необхідності розвитку неінвазивних методів діагностики, які забезпечують прийнятну достовірність та доступність.

4. Відсутність стандартизації медичних даних. Дані, які використовуються для діагностики, часто збираються за різними протоколами і форматами, що ускладнює їх інтерпретацію та інтеграцію в інформаційні системи.

5. Потреба у швидкому доступі до обробленої інформації. Традиційний процес діагностики потребує значних витрат часу на аналіз та оцінку результатів досліджень. Використання автоматизованих систем дозволяє прискорити цей процес, забезпечуючи лікарів швидким доступом до комплексної інформації про стан пацієнта.

Завдання, що мають бути вирішені для розробки ефективної ІСППР для діагностики ІХС:

1. Інтеграція та аналіз симптомів і діагностичних даних. Необхідно створити систему, яка здатна автоматично обробляти та інтегрувати дані з різних джерел, формуючи єдину картину стану пацієнта.

2. Розробка алгоритмів для попередньої діагностики. Алгоритми повинні враховувати всі фактори ризику та симптоми для формування попереднього діагнозу, що дозволить знизити кількість необхідних інвазивних досліджень.

3. Забезпечення безпеки та конфіденційності медичних даних. Оскільки система працює з чутливою інформацією про пацієнтів, важливо дотримуватися всіх норм щодо захисту даних.

4. Підвищення зручності користування системою. Інтерфейс ІСППР має бути інтуїтивно зрозумілим для лікарів, що дозволить їм швидко знаходити необхідну інформацію та приймати рішення без додаткових навчань.

5. Автоматизація рутинних завдань. Завдяки автоматизації збору та обробки даних система зменшує навантаження на медичний персонал, дозволяючи їм зосередитися на прийнятті рішень.

Аналіз предметної області показує, що для ефективної діагностики ІХС необхідна ІСППР, яка здатна інтегрувати різні джерела медичних даних, швидко їх обробляти та надавати рекомендації щодо діагнозу. Для цього важливо вирішити завдання автоматизації обробки симптомів і створити комплексний підхід, що поєднує традиційні методи діагностики з алгоритмами машинного навчання.

1.2 Аналіз методів для розпізнавання ішемічної хвороби серця

Успішна діагностика ішемічної хвороби серця (ІХС) потребує ефективного аналізу активності серця та динаміки його функцій під час різних фізичних та психоемоційних навантажень. Це вимагає інтеграції методів, які здатні розпізнавати різні аспекти активності серцевого м'яза, зміни артеріального тиску, ритму серця та інші фізіологічні параметри. Сучасні методи розпізнавання активності включають електрокардіографію (ЕКГ), ехокардіографію, холтерівський моніторинг, а також інноваційні технології на основі штучного інтелекту.

1. Електрокардіографія (ЕКГ). ЕКГ є основним неінвазивним методом для оцінки електричної активності серця. Метод дозволяє виявити аномалії ритму серця та зміни, які можуть свідчити про ішемічні процеси. Найбільш інформативною є ЕКГ з навантаженням, яка демонструє, як серце реагує на фізичну активність, виявляючи приховані ознаки ІХС. Однак метод обмежений тим, що він може не завжди виявляти всі форми ішемії, особливо на ранніх стадіях. На рисунку 1.1 зображено процедуру електрокардіографії.

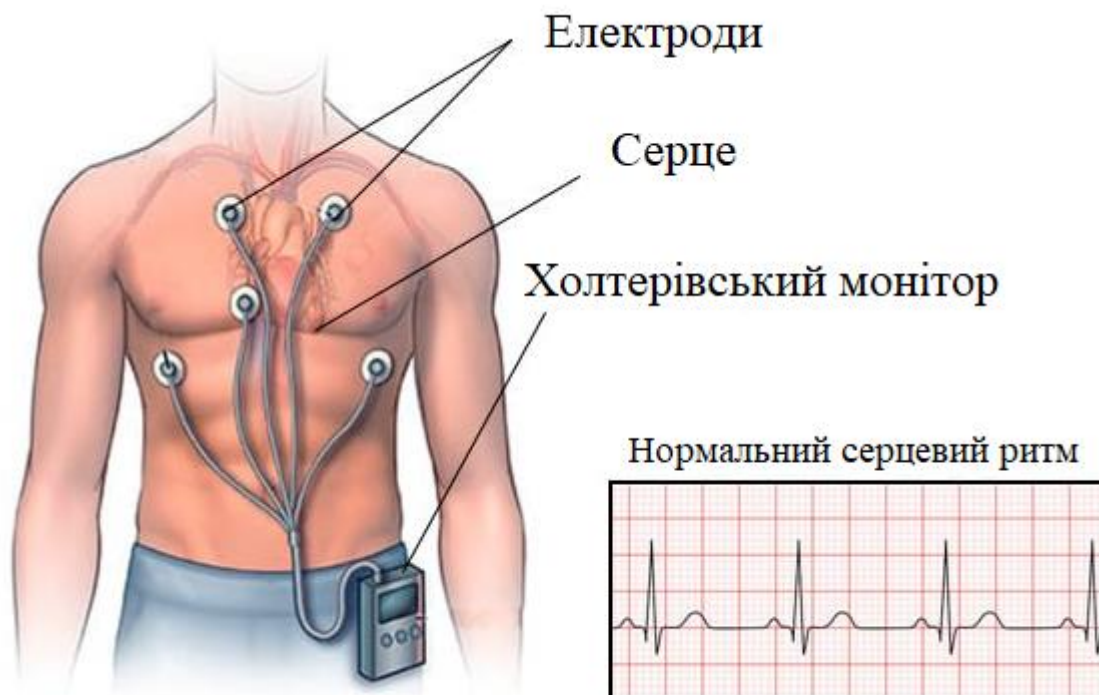


Рисунок 1.1 - Процедура електрокардіографії

2. Ехокардіографія. Цей метод дозволяє оцінити функціонування серця за допомогою ультразвукових хвиль. Ехокардіографія дає змогу отримати інформацію про стан серцевого м'яза, об'єм камер та товщину стінок, а також побачити функціонування клапанів. Вона ефективна для діагностики структурних змін, що можуть свідчити про розвиток ІХС, однак потребує високої кваліфікації медичного персоналу для інтерпретації результатів.

3. Холтерівський моніторинг. Холтерівське моніторування — це тривалий запис ЕКГ (протягом 24-48 годин), що дозволяє отримати повну картину змін ритму та електричної активності серця протягом доби. Це корисно для виявлення прихованої ішемії, особливо в нічний час або при стресових ситуаціях. Однак метод не завжди є зручним для пацієнтів і вимагає тривалої обробки даних.

4. Магнітно-резонансна томографія серця (МРТ серця). МРТ є високоінформативним методом для оцінки стану серцевого м'яза та кровообігу в ньому. Він дозволяє візуалізувати ішемічні ділянки та виявляти порушення

кровопостачання. Проте цей метод залишається дорогим і не завжди доступним для пацієнтів через високі витрати та технічні вимоги.

5. Використання методів штучного інтелекту та машинного навчання. Останнім часом у діагностиці ІХС дедалі частіше застосовують алгоритми штучного інтелекту. Вони дозволяють автоматизувати процес розпізнавання активності серця на основі даних ЕКГ, ехокардіографії або МРТ. Моделі машинного навчання можуть аналізувати великі обсяги даних, виявляти закономірності та здійснювати попереднє діагностування на основі ознак, які можуть бути неочевидними для лікаря. Алгоритми, зокрема глибокі нейронні мережі, мають здатність адаптуватися та вдосконалюватися, підвищуючи достовірність розпізнавання та прогнозування розвитку ІХС.

Кожен із розглянутих методів має свої переваги та обмеження у діагностиці ІХС. ЕКГ є швидким та доступним, але має обмежену достовірність у певних ситуаціях. Ехокардіографія дозволяє отримати детальну інформацію про анатомічний стан серця, але потребує високої кваліфікації персоналу. Холтерівський моніторинг забезпечує тривалий запис даних, але вимагає додаткового часу для обробки. МРТ є надзвичайно інформативним, але не завжди доступний. Методи штучного інтелекту, у свою чергу, забезпечують високу достовірність і швидкість обробки, але потребують великих обсягів даних для навчання та значних обчислювальних ресурсів.

Оптимальний підхід до розпізнавання активності для діагностики ІХС часто включає поєднання кількох методів. Це дозволяє отримати комплексну інформацію про стан серця та зменшити ризик пропуску прихованих ознак ішемії. Крім того, застосування штучного інтелекту для автоматизації аналізу даних з ЕКГ, ехокардіографії або холтерівського моніторингу може значно підвищити ефективність діагностики, особливо у великих медичних закладах з великою кількістю пацієнтів.

Таким чином, аналіз методів розпізнавання активності серця для діагностики ІХС показує, що комплексний підхід, який поєднує традиційні методи та технології штучного інтелекту, є найбільш перспективним. Це

дозволяє підвищити достовірність і доступність діагностичних процесів, що особливо важливо для своєчасного виявлення та лікування ішемічної хвороби серця.

1.3 Методи м'яких обчислень у системах діагностики

Методи м'яких обчислень, які включають ряд методологій для наближеного пошуку рішень, широко застосовуються в реальних задачах, зокрема у медичній діагностиці. Використання таких методів для розробки діагностичних систем мотивується їхньою здатністю обробляти невизначеність, що часто супроводжує медичні діагнози.

Штучні нейронні мережі (ШНМ) та метод опорних векторів (МОВ) є найпоширенішими методами м'яких обчислень у розробці систем для діагностики ІХС. Ці системи демонструють обнадійливі результати завдяки відносно високій достовірності класифікації, що робить їх надійними інструментами для підтримки прийняття рішень при виявленні ІХС. Однак, незважаючи на їхню високу достовірність, ШНМ і МОВ часто критикують через низьку прозорість, оскільки вони функціонують як «чорні ящики», що не надають користувачеві зрозумілої інформації про процес ухвалення рішень.

Останнім часом проблема прозорості, яка відображає здатність системи зрозуміло представити свою логіку, привертає все більше уваги. Прозорість вважається важливим аспектом у медичних діагностичних системах, оскільки вона дозволяє лікарям аналізувати і приймати або відкидати рішення, запропоновані системою. Це є особливо важливим для виявлення факторів ризику ІХС на ранніх і потенційно зворотних стадіях.

На відміну від ШНМ та МОВ, системи на основі нечітких правил забезпечують прозоре та інтерпретоване представлення знань у вигляді лінгвістичних умовних тверджень типу «якщо умова(и), то дія(ї)». Такий формат є близьким до природної мови і зручним для аналізу та інтерпретації людиною. Крім того, НСНП використовують механізм міркувань, заснований

на нечіткій логіці, що імітує людський спосіб мислення. Ці особливості роблять НСНП перспективним методом для розробки систем діагностики ІХС.

Нечіткі правила можуть генеруватися або на основі експертних знань, або автоматично з даних. У першому випадку параметри, такі як межі інтервалів і функції належності, визначаються вручну експертами. У другому випадку їх значення автоматично встановлюються на основі набору репрезентативних прикладів з використанням методів навчання.

Наразі в діагностиці ІХС переважають методи автоматичної генерації правил, оскільки вони дозволяють використовувати історичні дані пацієнтів і нові алгоритми машинного навчання, що підвищує достовірність діагностики. Проте, такі методи зазвичай нехтують прозорістю, і нечіткі системи часто працюють як «чорні ящики». При розробці діагностичних систем на основі нечітких правил важливо досягти балансу між точністю та прозорістю.

Генетичні алгоритми є одним із найуспішніших підходів до генерації прозорих нечітких систем, які балансують між точністю та прозорістю. Однак вони можуть бути обчислювально витратними, особливо для великих обсягів даних, тому бажаним є попередній відбір релевантних ознак.

Деякі сучасні підходи також намагаються забезпечити прозорість, надаючи додаткову інформацію про правила, наприклад, ступінь їхньої достовірності та підтримки в наборі даних, що допомагає особам, які ухвалюють рішення, краще оцінити значимість кожного правила.

1.4 Опис відомих наукових методів і систем

Сучасні наукові дослідження в області діагностики та моніторингу ішемічної хвороби серця (ІХС) активно використовують новітні технології, такі як машинне навчання, глибоке навчання, аналітику великих даних. Багато досліджень спрямовані на розробку інтелектуальних систем, які дозволяють проводити раннє виявлення та прогнозування захворювань, мінімізуючи ризик ускладнень і покращуючи якість життя пацієнтів. У зв'язку з цим розглядаються

різноманітні підходи, що включають аналіз біомедичних сигналів, виявлення ключових клінічних ознак, обробку даних про активність та поведінку пацієнтів.

Потреба в таких системах зростає через значне поширення ІХС у світі, яка залишається однією з основних причин смертності та інвалідизації серед дорослого населення. Створення ефективної інформаційної системи для підтримки прийняття рішень у діагностиці ІХС є актуальним завданням для медицини, оскільки це дозволяє значно підвищити достовірність діагностики, зменшити ризик хибнопозитивних або хибнонегативних результатів, а також знизити витрати на лікування за рахунок своєчасного виявлення захворювання. У сучасних умовах існує безліч підходів до створення таких систем, які можуть інтегруватися з клінічними базами даних, мобільними пристроями та спеціалізованими медичними датчиками.

На сьогоднішній день ряд досліджень показали, що застосування методів машинного навчання для аналізу медичних даних значно підвищує достовірність прогнозування захворювань. При цьому використовуються як традиційні методи, такі як логістична регресія і дерева рішень, так і сучасні нейронні мережі, особливо глибокі нейронні архітектури та рекурентні мережі з короткостроковою та довгостроковою пам'яттю (LSTM). Крім того, деякі системи впроваджують елементи штучного інтелекту для автоматичного виявлення та аналізу фізіологічних сигналів пацієнтів, що дозволяє створювати індивідуалізовані рішення.

У роботі [2] запропонована система, що базується на глибоких нейронних мережах для аналізу ЕКГ-сигналів пацієнтів з ішемічною хворобою серця. Дана система використовує згорткові нейронні мережі для виділення значущих ознак, що сприяє ранньому виявленню ІХС, та демонструє високу достовірність у порівнянні з традиційними методами. У дослідженні [3] представлено методи машинного навчання для класифікації пацієнтів з ризиком ІХС на основі клінічних даних.

У роботі [4] розглянута гібридна модель, яка поєднує логістичну регресію з підтримкою векторних машин. Модель протестована на базі даних Framingham Heart Study і продемонструвала зниження рівня хибнопозитивних результатів, що покращує достовірність прогнозування ризику серцево-судинних захворювань. Інше дослідження [5] описує інтегровану систему для аналізу ангіографічних зображень коронарних судин із застосуванням методів глибокого навчання, що дозволяє автоматично визначати ступінь ураження судин і прискорює діагностичний процес.

У дослідженні [6] розроблено систему, яка аналізує комбінації біомаркерів крові, використовуючи алгоритми класифікації, такі як К-найближчих сусіди (KNN) та дерева рішень, що дозволило досягти високої ефективності в ранньому виявленні ІХС. У роботі [7] запропонована модель прогнозування ризику ІХС, заснована на нейронних мережах з короткостроковою та довгостроковою пам'яттю (LSTM), яка використовує часові ряди для аналізу показників артеріального тиску, рівня глюкози та серцевої активності, показуючи високу достовірність у довгостроковому прогнозуванні.

Дослідження [8] присвячено використанню логістичної регресії та нейронних мереж для аналізу факторів ризику ІХС, таких як рівень холестерину, артеріальний тиск та індекс маси тіла. Результати показали, що комплексний підхід до аналізу клінічних даних дозволяє підвищити достовірність прогнозування ризиків. У роботі [9] створена платформа для підтримки прийняття рішень на основі мобільних пристроїв, яка інтегрує аналітичні інструменти, що спрощує її використання в клінічній практиці.

У дослідженні [10] представлена адаптивна система, що враховує індивідуальні особливості пацієнтів, дозволяючи підвищити достовірність діагностики в складних випадках. Нарешті, у роботі [11] розроблено платформу для прогнозування ризику ІХС на основі глибокого навчання, яка об'єднує дані ЕКГ і клінічні показники, що значно підвищує ефективність передбачення захворювання. Всі ці роботи демонструють значний потенціал інноваційних

технологій для покращення достовірності та ефективності діагностики ІХС, що особливо важливо для зниження смертності від серцево-судинних захворювань.

З наведеного огляду літератури можна зробити висновок, що розвиток подібних систем сприяє поліпшенню якості медичних послуг та забезпечує більш високу ефективність діагностики ІХС. Незважаючи на значний прогрес у цій галузі, багато наукових проблем залишаються відкритими. Серед них – оптимізація обробки великих обсягів даних, підвищення адаптивності моделей до індивідуальних особливостей пацієнтів, забезпечення конфіденційності та захисту персональних даних, що є особливо важливим для медичних застосувань.

Проведений огляд наукових робіт дозволяє зробити висновок про необхідність подальших досліджень у напрямку створення гнучких та адаптивних інформаційних систем, що забезпечують високу достовірність діагностики ІХС в умовах реального використання. Особливий інтерес представляють інтегровані системи, які можуть працювати з різними джерелами даних, такими як ЕКГ-сигнали, біомаркери крові, ангиографічні зображення, а також дані з мобільних пристроїв пацієнтів. Такий підхід дозволяє створювати багатокomпонентні рішення, здатні враховувати широкий спектр факторів ризику та забезпечувати більш персоналізоване прогнозування.

Далі у роботі буде розглянуто розробку прототипу подібної системи, а також проведено серію експериментів з метою оцінки її ефективності. Особливу увагу буде приділено дослідженню взаємодії між різними модулями системи, перевірці достовірності та надійності її роботи, а також виявленню можливих шляхів оптимізації для забезпечення кращих результатів у діагностиці ІХС.

1.5 Опис даних та методів їх обробки

Набір даних, використаний у цьому дослідженні, було отримано з UCI Repository of Machine Learning Databases (Репозиторій баз даних машинного

навчання). Він був зібраний з Клівлендської клінічної фундації, Клівленд, штат Огайо, і наданий Робертом Детрано, доктором медицини, доктором філософії Медичного центру V.A., Лонг-Біч, Каліфорнія. Спочатку дані містили 76 атрибутів, з яких для використання було обрано лише 13 атрибутів. Відібрані ознаки представляють результати клінічних і неінвазивних досліджень 303 пацієнтів, які проходили ангиографію. Вилучивши випадки, що містили відсутні значення, в дослідженні було розглянуто 270 випадків, з яких 120 випадків були ідентифіковані як пацієнти з ІХС, а 150 випадків - як пацієнти без ІХС.

У таблиці 1.1 наведено короткий опис атрибутів даних Клівленда та їхніх відповідних значень.

Таблиця 1.1 – Опис набору даних Клівлендського дослідження серцево-судинних захворювань

Атрибути	Опис	Тип	Значення
Вік	Вік	Integer	[29, 77]
Стать	Стать	Integer	1 = чоловік; 0 = жінка;
ТГБ	Тип болю в грудях	Integer	1 = типова стенокардія; 2 = атипова стенокардія; 3 = неангінозний біль; 4 = безсимптомний;
АТС	Артеріальний тиск у стані спокою (у мм рт. ст. при надходженні до лікарні)	Integer	[94, 200]
РХС	Рівень холестерину в сироватці крові (мг/дл)	Integer	[126, 564]
РГКН	Рівень глюкози в крові натще (>120 мг/дл)	Integer	1 = так; 0 = ні;
РЕКС	Результати електрокардіограми в стані	Integer	[0, 2]

	спокою (значення 0, 1, 2)		
--	---------------------------	--	--

Кінець таблиці 1.1

Атрибути	Опис	Тип	Значення
МЧСС	Максимальна досягнута частота серцевих скорочень	Integer	[71, 202]
СФН	Стенокардія, викликана фізичним навантаженням	Integer	1 = так; 0 = ні
ДСТ	Депресія ST сегмента, викликана фізичним навантаженням, у порівнянні зі станом спокою	Real	[0.00, 62.00]
НПСТ	Нахил пікового сегмента ST під час фізичного навантаження	Integer	1 = нахил вгору; 2 = плоский; 3 = нахил вниз;
КМС	Кількість магістральних судин (0–3), візуалізованих за допомогою флуороскопії	Integer	[0, 3]
Таласемія	Таласемія	Integer	3 = норма; 6 = фіксований дефект; 7 = оборотний дефект;
ІХС	Діагностика ішемічної хвороби серця	Integer	0 = відсутня; 1 = присутня;

1.6 Постановка задачі дослідження

Постановка задачі є одним із ключових етапів наукового дослідження, оскільки саме від чіткого формулювання завдання залежить успішність розробки та впровадження інформаційної системи підтримки прийняття рішень (ІСППР) для діагностики ішемічної хвороби серця (ІХС). Дослідження в цій

галузі є особливо актуальним, оскільки ішемічна хвороба серця є однією з провідних причин смертності у світі, а точна діагностика та раннє виявлення патології мають вирішальне значення для ефективного лікування та запобігання розвитку ускладнень.

Основна мета дослідження полягає в розробці, реалізації та оцінці ефективності інформаційної системи, яка здатна допомогти медичному персоналу в процесі діагностики ІХС. Ця система повинна автоматизувати процес обробки медичних даних, використовуючи методи машинного навчання та штучного інтелекту для аналізу результатів діагностичних тестів, таких як електрокардіограма (ЕКГ), лабораторні аналізи, а також інші медичні показники. Завданням дослідження є не лише розробка системи, але й оцінка її здатності покращити достовірність і швидкість діагностики.

Ішемічна хвороба серця є серйозною медичною проблемою, яка викликає зниження якості життя пацієнтів та високу смертність. Вона часто залишається нерозпізнаною на ранніх стадіях через нечіткі симптоми або недостатню увагу до результатів діагностичних досліджень. Сучасні методи діагностики ІХС, хоча й ефективні, можуть вимагати значних витрат часу та ресурсів, що може призводити до затримок у прийнятті рішень щодо лікування.

Таким чином, постає завдання створити таку систему, яка може в реальному часі аналізувати різні джерела медичних даних та допомагати лікарям у постановці точних діагнозів. Запропонована інформаційна система повинна мати здатність інтегрувати дані з різних джерел, таких як результати лабораторних аналізів, дані обстеження пацієнта, а також зібрані сигнали ЕКГ. Окрім того, система повинна виконувати класифікацію та аналіз на основі методів машинного навчання, щоб забезпечити підтримку прийняття рішень лікарем.

Основною метою дослідження є створення інформаційної системи, яка сприятиме покращенню діагностики ІХС через автоматизацію процесу обробки медичних даних та надання рекомендацій лікарям на основі аналізу великого масиву інформації. Для досягнення цієї мети було поставлено такі завдання:

1. Аналіз існуючих методів діагностики ІХС та їхніх недоліків. На цьому етапі необхідно вивчити сучасні методи та підходи до діагностики ІХС, включаючи інструментальні методи, як-от ЕКГ та ангіографія, а також використання лабораторних показників. Крім того, важливо визначити основні недоліки існуючих рішень, такі як затримки у проведенні досліджень, недостатня достовірність та суб'єктивний фактор у діагностиці.

2. Розробка моделі інформаційної системи для діагностики ІХС. Це завдання передбачає проектування архітектури системи, яка інтегрує в собі кілька модулів: модуль збору та інтеграції даних, модуль обробки сигналів, модуль машинного навчання для класифікації стану пацієнта, а також модуль підтримки прийняття рішень та інтерфейс користувача. Модель повинна забезпечувати швидку обробку великого обсягу даних та видавати точні результати.

3. Реалізація та тестування інформаційної системи. Після розробки архітектури необхідно створити прототип системи, який буде перевірений на реальних даних. Система має здатність обробляти різні типи медичних даних, виконувати їхню класифікацію за допомогою алгоритмів машинного навчання та видавати рекомендації щодо діагностики та лікування ІХС. На цьому етапі також важливо провести тестування системи на предмет її продуктивності, достовірності та надійності.

4. Порівняльний аналіз результатів роботи системи з існуючими методами. Щоб оцінити ефективність розробленої системи, необхідно провести порівняльні експерименти, у яких система буде використана для діагностики ІХС на різних наборах даних. Результати роботи системи повинні бути порівняні з традиційними методами діагностики, що дозволить оцінити її переваги та недоліки.

5. Оцінка впливу інформаційної системи на клінічний процес діагностики ІХС. Окрім технічної оцінки системи, необхідно дослідити, як вона впливає на роботу лікарів та процес прийняття рішень у реальних клінічних умовах. Чи допомагає система скоротити час діагностики? Чи підвищує вона

достовірність постановки діагнозу? Чи зменшує кількість помилкових рішень? Відповіді на ці питання допоможуть оцінити реальну корисність системи в медичній практиці.

В результаті проведеного дослідження очікується створення прототипу інформаційної системи, здатної автоматизувати процес діагностики ІХС та підтримувати медичних працівників у прийнятті рішень. Система повинна продемонструвати високу достовірність у класифікації стану пацієнтів, швидкість обробки великих обсягів даних, а також забезпечити надійний захист інформації. Окрім того, система має бути легкою у використанні для лікарів, що забезпечить її швидку адаптацію у клінічній практиці.

Дослідження також дозволить отримати важливу інформацію щодо практичного використання систем штучного інтелекту та машинного навчання у сфері медичної діагностики. Це відкриє нові можливості для подальших досліджень у галузі створення інтелектуальних систем підтримки медичних рішень і допоможе підвищити якість надання медичних послуг.

1.7 Висновки

Аналіз наукової літератури, проведений у цьому розділі, свідчить про важливість використання сучасних підходів для діагностики ішемічної хвороби серця (ІХС). Діагностика ІХС є складним завданням через широкий спектр можливих симптомів, які часто можуть бути невиразними або маскуватися під інші захворювання. У зв'язку з цим, традиційні методи, засновані на аналізі клінічних даних, мають певні обмеження і часто потребують додаткових інструментів для підвищення достовірності.

В останні десятиліття методи машинного навчання та м'яких обчислень, включно з нейронними мережами, методом опорних векторів (МОВ) і нечіткими системами, довели свою ефективність для покращення процесу діагностики. Їх застосування в медичній діагностиці дозволяє обробляти великі обсяги даних, автоматично визначати приховані закономірності та підвищувати

достовірність постановки діагнозу. У той же час, використання цих методів приносить нові виклики, такі як труднощі в інтерпретованості отриманих результатів, що може бути перешкодою для медичних працівників при використанні таких моделей на практиці.

Серед основних переваг методів машинного навчання можна виділити їхню здатність адаптуватися до нових даних і покращувати свою ефективність із часом. Наприклад, нейронні мережі здатні до самонавчання, що дозволяє їм знаходити більш точні діагностичні рішення з кожним новим аналізованим випадком. Метод опорних векторів забезпечує високий рівень достовірності за рахунок побудови оптимальних меж між класами даних, що також може бути корисним у діагностиці ІХС. Нечіткі системи, своєю чергою, дають можливість оперувати з невизначеністю в даних, що часто зустрічається в медичних діагнозах.

Незважаючи на високу ефективність, ці методи мають певні обмеження. Одним із них є складність у розумінні та інтерпретації результатів, що часто викликає труднощі при їхньому впровадженні в медичні системи підтримки прийняття рішень (СППР). Прозорість моделі є важливим критерієм для таких систем, оскільки лікарі повинні розуміти логіку постановки діагнозу, щоб повноцінно довіряти автоматизованим рішенням і приймати обґрунтовані клінічні рішення. Тому в наукових дослідженнях все частіше піднімається питання створення прозорих діагностичних моделей, які б поєднували достовірність з інтерпретованістю.

Таким чином, узагальнюючи результати огляду літератури, можна зробити висновок, що сучасні методи машинного навчання та нечіткі системи пропонують перспективні можливості для підвищення достовірності та надійності діагностики ішемічної хвороби серця.

2 МОДЕЛЬ ПРОЦЕСУ ДІАГНОСТИКИ ШЕМІЧНОЇ ХВОРОБИ СЕРЦЯ

2.1 Етапи розробки та методологія

У цьому дослідженні розроблено прозору систему на основі нечітких правил для діагностики ІХС. Поняття прозорості розширено, щоб включити не тільки традиційні показники, такі як кількість правил і нечітких множин, але й ступінь достовірності та підтримки кожного правила. Розробка нашої системи, заснованої на нечітких правилах, включає чотири основні етапи. На першому етапі, який є етапом попередньої обробки, спрямованим на зменшення складності набору даних, застосовується метод відбору ознак для вибору найбільш дискримінаційної підмножини ознак. Після генерації всіх можливих нечітких правил на другому кроці, на третьому кроці використовуються багатоцільові генетичні алгоритми для отримання підмножини невеликої кількості правил з найвищою класифікаційною здатністю. Результатом цього кроку є набір нечітких правил, які включають в себе, окрім міток класу, ступінь їх визначеності та достовірності.

Для подальшого підвищення достовірності отриманих нечітких правил, на четвертому кроці вводиться новий метод, в якому рішення або класифікація, зроблена правилом, ступінь достовірності якого є меншою за порогове значення, підтримується ансамблем вибраних класифікаторів для отримання більш надійного рішення. Ця стратегія ансамблю класифікаторів (САК) натхненна поширеною медичною практикою, коли в складних випадках з низькою діагностичною достовірністю необхідно враховувати думку більш ніж одного експерта. Результати, отримані за допомогою НСНП, оцінюються і порівнюються з іншими еталонними методами.

2.2 Архітектура інформаційної системи для діагностики ІХС

Архітектура інформаційної системи підтримки прийняття рішень (ІСППР) для діагностики ішемічної хвороби серця (ІХС) є ключовим компонентом для забезпечення ефективної та точної обробки медичних даних. Така система повинна забезпечувати комплексну інтеграцію даних із різних джерел, обробку цих даних із використанням передових методів машинного навчання, а також підтримку медичних рішень на основі результатів аналізу. Крім того, особлива увага приділяється захисту чутливих медичних даних, які використовуються у процесі діагностики та прийняття рішень.

Архітектура інформаційної системи для діагностики ІХС є модульною та складається з кількох основних компонентів, кожен із яких виконує специфічну функцію у процесі обробки інформації. Нижче наведено детальний опис кожного з цих модулів та їх взаємодії, а також пояснення ролі кожного з них у загальній системі.

Основні компоненти архітектури ІСППР:

1. Модуль збору та інтеграції даних. Цей модуль відповідає за отримання та об'єднання даних із різних джерел. Дані надходять від медичних записів пацієнтів, діагностичних пристроїв (наприклад, ЕКГ), лабораторних тестів та інших зовнішніх систем. Інтеграція даних є важливим етапом, оскільки вона дозволяє зібрати всі необхідні відомості в єдиний формат для подальшої обробки. Збір даних здійснюється автоматично, що дозволяє мінімізувати втрати інформації та уникнути помилок, пов'язаних із ручним введенням.

2. Модуль попередньої обробки даних. Після отримання даних вони часто потребують очищення від шуму, усунення аномалій, заповнення пропущених значень або перетворення до єдиного формату. У цьому модулі відбувається нормалізація даних, що дозволяє забезпечити їхню подальшу коректну обробку в наступних етапах. Попередня обробка відіграє вирішальну

роль, оскільки наявність помилок або шуму може значно вплинути на достовірність прогнозу.

3. Модуль аналізу та обробки сигналів. Один із ключових етапів обробки медичних даних для діагностики ІХС – це аналіз сигналів, таких як ЕКГ. Цей модуль відповідає за обробку сигналів для виявлення критичних ознак ішемічної хвороби. Використовуються різноманітні алгоритми обробки сигналів, які можуть виявити патерни, що свідчать про порушення серцевого ритму або інші проблеми, характерні для ІХС.

4. Модуль машинного навчання та класифікації. Основна задача цього модуля – навчання та застосування моделей машинного навчання для класифікації стану пацієнта на основі оброблених даних. Тут використовуються різні алгоритми, як-от дерева рішень, нейронні мережі, методи підтримки векторів, які дозволяють системі з високою точністю визначати наявність або ризику розвитку ІХС. Цей модуль має можливість постійного навчання на нових даних, що підвищує достовірність і надійність діагностики.

5. Модуль підтримки прийняття рішень. Цей модуль на основі результатів роботи попередніх етапів генерує рекомендації для лікаря. Система надає поради щодо подальших діагностичних тестів або лікувальних заходів. Такі рекомендації допомагають лікарю швидко приймати рішення щодо лікування пацієнта, ґрунтуючись на наукових даних і попередніх випадках.

6. Модуль інтерфейсу користувача. Модуль інтерфейсу користувача відповідає за забезпечення взаємодії медичного персоналу із системою. Інтерфейс повинен бути інтуїтивно зрозумілим і зручним, дозволяючи лікарю швидко переглядати результати діагностики та рекомендації. Інтерфейс також забезпечує зворотний зв'язок із системою, дозволяючи лікарям вводити додаткові дані для подальшої обробки.

7. Модуль безпеки та захисту даних. Захист даних є критично важливим компонентом архітектури, оскільки система працює з чутливими медичними даними. Модуль забезпечує захист від несанкціонованого доступу, шифрування даних, контроль за аутентифікацією користувачів та ведення

журналів доступу. Безпека є ключовим елементом для підтримки відповідності законодавчим вимогам та забезпечення конфіденційності пацієнтів.

Модульна архітектура системи забезпечує чітку та ефективну взаємодію між усіма компонентами. Дані надходять до системи через модуль збору та інтеграції даних, де вони проходять попередню обробку. Далі оброблені дані передаються до модуля аналізу сигналів і модулів машинного навчання для проведення класифікації. Результати класифікації та рекомендації щодо прийняття рішень передаються лікарю через інтерфейс користувача. Модуль безпеки контролює всі етапи обробки, забезпечуючи конфіденційність і безпеку даних на кожному етапі.

Для кращого розуміння архітектури системи надається рисунок 2.1, який візуально показує взаємодію основних модулів та етапи обробки даних у процесі діагностики ІХС.

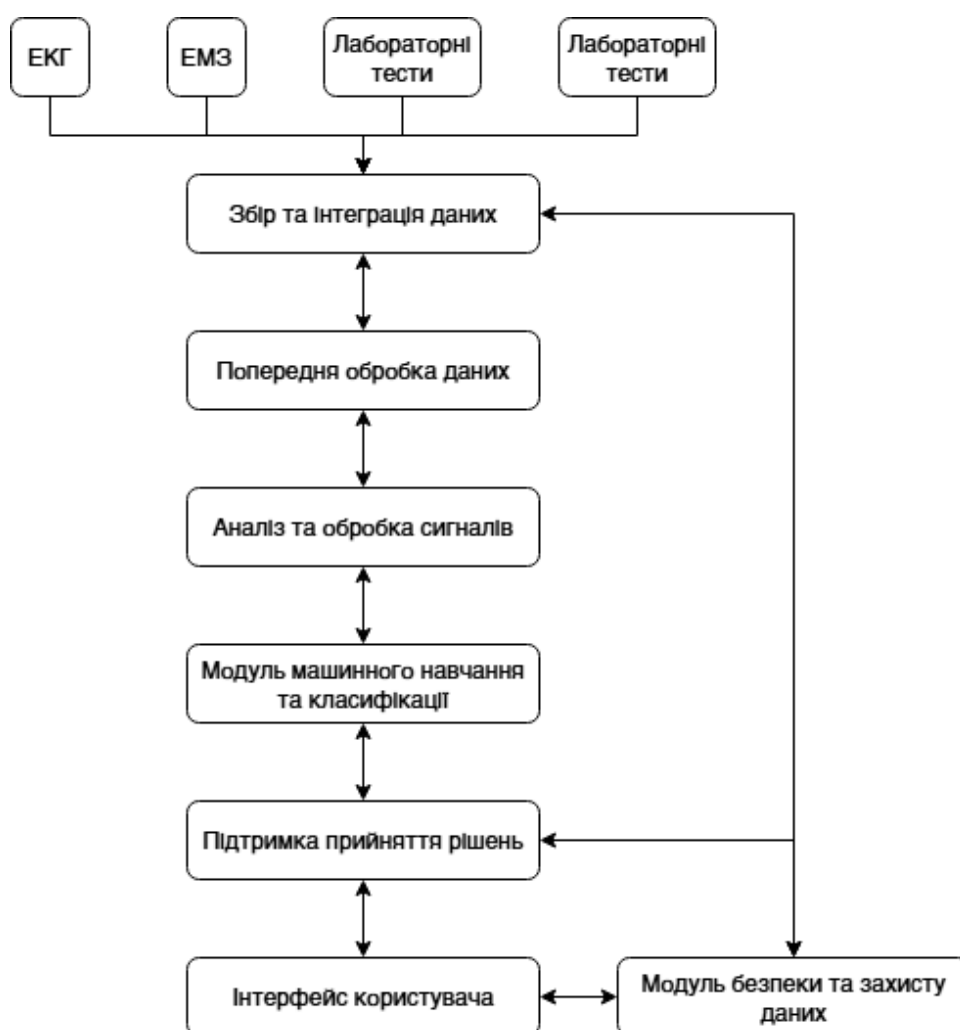


Рисунок 2.1 - Діаграма взаємодія та етапи обробки даних у процесі діагностики ІХС

2.3 Модель процесу діагностики ішемічної хвороби серця

Модель процесу діагностики ішемічної хвороби серця (ІХС) є багатоступеневим підходом, який передбачає послідовну реалізацію кількох важливих етапів для забезпечення максимальної достовірності діагностики. Процес починається зі збору даних, що є основою для всіх подальших кроків. На цьому етапі відбувається збір клінічних і лабораторних показників пацієнта, таких як базові фізіологічні дані (вік, стать, артеріальний тиск) і специфічні медичні обстеження (електрокардіограма, ехокардіографія, аналізи крові тощо). Всі ці дані створюють базу для подальшого аналізу, забезпечуючи надійну основу для виявлення потенційних ознак ішемії.

Після збору даних здійснюється етап попередньої обробки, метою якого є очищення та нормалізація отриманої інформації, а також відбір найважливіших ознак, які мають найбільший вплив на діагностику ІХС. Використання методів відбору ознак дозволяє зменшити обсяг даних, фокусуючись на ключових параметрах, що сприяє підвищенню ефективності та швидкості роботи моделі. Цей етап є критично важливим для запобігання перенаванчання моделі та підвищення достовірності її прогнозів.

Наступним етапом є побудова діагностичної моделі. На цьому кроці застосовуються методи штучного інтелекту, зокрема, нейронні мережі, методи опорних векторів та нечіткі системи. Використання нечітких правил у моделі дає можливість представити медичні діагностичні критерії у формі, що є зрозумілою для лікарів і дозволяє створити прозору систему для підтримки діагностичних рішень. Для підвищення достовірності моделі застосовуються багатоцільові генетичні алгоритми, які допомагають обрати підмножину правил з найвищою класифікаційною здатністю, забезпечуючи баланс між точністю і прозорістю.

Ключовим етапом є оцінка достовірності отриманих результатів. Прозорість моделі дозволяє лікарю легко інтерпретувати діагностичні рішення, що є важливим для медичної практики. На цьому етапі визначається ступінь достовірності кожного правила, що дозволяє медичним працівникам оцінити, наскільки впевнено модель робить висновки. Важливим аспектом є використання методу ансамблю класифікаторів для отримання більш надійного результату у випадках низької достовірності, що імітує медичний консилиум.

Після завершення діагностичного процесу результати інтегруються у систему підтримки прийняття рішень (СППР), яка надає лікареві рекомендації щодо подальших дій. Висновки моделі відображають не лише ймовірність наявності ІХС у пацієнта, але й ключові ознаки, які вплинули на діагноз. Це дозволяє лікареві не лише отримати результат, але й зрозуміти підґрунтя рішення.

Для підтвердження ефективності розробленої моделі проводиться порівняння результатів з іншими методами діагностики, що дозволяє оцінити її достовірність, прозорість і достовірність у порівнянні з альтернативними підходами. Це дає можливість виявити, наскільки розроблена модель є конкурентоспроможною і ефективною для діагностики ІХС.

На рисунку 2.2 відображені основні етапи моделі діагностики у вигляді блок-схеми.



Рисунок 2.2 – Модель процесу діагностики ішемічної хвороби серця

2.4 Моделі машинного навчання для діагностики ішемічної хвороби серця

Сучасні системи діагностики захворювань, зокрема ішемічної хвороби серця, все частіше базуються на застосуванні методів машинного навчання. Завдяки можливості виявляти приховані закономірності в складних наборах медичних даних, ці методи дозволяють підвищити достовірність та швидкість діагностики, що, своєю чергою, позитивно впливає на результати лікування. У цьому підрозділі буде розглянуто основні методи машинного навчання, що використовуються в системах діагностики ІХС, такі як штучні нейронні мережі, метод опорних векторів, дерева рішень та інші підходи. Кожен з цих методів має свої особливості, переваги та недоліки, які впливають на їхню ефективність в умовах медичної діагностики.

Штучні нейронні мережі (ШНМ) є одними з найпоширеніших методів у медичній діагностиці. Вони складаються з шарів штучних нейронів, які пов'язані між собою та можуть "навчатися" на великій кількості даних. ШНМ здатні виконувати високоточну класифікацію та прогнозування, оскільки кожен шар у мережі витягує певні ознаки або характеристики з набору даних, що

надходить з попереднього шару. В умовах діагностики ІХС це означає, що ШНМ можуть виявляти приховані взаємозв'язки між клінічними ознаками захворювання та його ймовірністю, що не завжди очевидно для людського спостереження.

Проте основним недоліком ШНМ є їхня недостатня прозорість, що часто описують як проблему "чорного ящика". Це означає, що, незважаючи на високу достовірність класифікації, неможливо точно зрозуміти, як мережа приймає певні рішення. У контексті медичної діагностики така особливість може ускладнювати процес прийняття рішення лікарем, оскільки не завжди зрозуміло, на основі яких характеристик мережа визначає діагноз.

Метод опорних векторів (МОВ) є ще одним потужним інструментом для розпізнавання та класифікації медичних даних. Суть МОВ полягає в тому, що він створює гіперплощину, яка максимально розділяє дані між класами. У діагностиці ІХС це дозволяє відокремити здорових пацієнтів від тих, що мають захворювання, на основі заданих ознак. МОВ також здатен працювати з даними високої розмірності, що є актуальним для медичних даних, де зазвичай багато змінних та параметрів.

Основна перевага МОВ полягає в його високій достовірності навіть за наявності невеликої кількості навчальних даних. Крім того, цей метод може забезпечувати досить зрозумілу інтерпретацію в контексті медичної діагностики, оскільки він чітко вказує на ті ознаки, які найбільшою мірою впливають на прийняття рішення. Однак, подібно до ШНМ, МОВ також має деякі недоліки, серед яких - висока обчислювальна вартість при роботі з великими обсягами даних та можливі складнощі з інтерпретацією в разі складних розподілів ознак.

Дерева рішень є одним із найзрозуміліших методів машинного навчання, оскільки вони представляють процес прийняття рішення у вигляді дерева, де кожне розгалуження відповідає певній умові або ознаці. Цей підхід дозволяє створювати прозорі системи підтримки прийняття рішень, де чітко видно, на основі яких критеріїв приймається діагностичне рішення. В умовах діагностики

ІХС дерево рішень може допомогти виявити найбільш значущі ознаки, які визначають наявність або відсутність захворювання.

Основною перевагою дерев рішень є їхня інтерпретованість: лікар може легко простежити весь шлях прийняття рішення. Однак, дерева рішень схильні до проблеми перенавчання, коли модель добре працює на навчальних даних, але показує слабші результати на нових. Крім того, складні медичні дані з численними ознаками можуть призвести до створення дуже великих дерев, що також знижує їхню інтерпретованість.

Ансамблеві методи передбачають поєднання кількох моделей для отримання більш стабільного та точного результату. До ансамблевих методів належать такі підходи, як підсилення, метод забору з поверненням та випадковий ліс. Кожен з цих методів використовує сукупність моделей, що працюють разом, щоб мінімізувати похибки, які можуть виникнути при використанні однієї моделі.

Для діагностики ІХС ансамблеві методи можуть бути особливо корисними, оскільки вони дозволяють знизити ризик хибнопозитивних або хибнонегативних результатів. Зокрема, випадковий ліс, який використовує безліч дерев рішень, є досить популярним методом для медичної діагностики через свою високу достовірність і стабільність. Крім того, ансамблеві методи можуть допомогти врахувати різні аспекти даних про пацієнта, що підвищує надійність моделі.

Застосування методів відбору ознак є важливим етапом у процесі побудови моделі діагностики ІХС, оскільки не всі зібрані дані можуть мати вирішальне значення для прийняття рішення. Відбір ознак дозволяє зосередитися на найбільш значущих змінних, що знижує складність моделі і підвищує її продуктивність. У цьому контексті методи відбору ознак, як-от послідовний прямий пошук або інші евристичні методи, допомагають зменшити кількість вхідних даних, зберігаючи при цьому їхню інформативність.

2.4.1 Математичні основи методів машинного навчання

У цьому підрозділі розглядаються основні математичні основи методів машинного навчання, які використовуються для побудови моделей для діагностики ішемічної хвороби серця. Нижче перераховані основні методи, які застосовуються в цій роботі для побудови ефективних моделей класифікації:

1. Метод опорних векторів. Метод опорних векторів базується на ідеї пошуку гіперплощини, яка максимально розділяє різні класи в багатовимірному просторі ознак. Основне рівняння МОВ, яке визначає гіперплощину, можна записати як:

$$w \cdot x + b = 0 \quad (1)$$

де w — вектор ваг, x — вектор вхідних ознак, а b — зміщення. Основна задача МОВ полягає у мінімізації функції втрат, яка враховує як відстань до гіперплощини, так і похибку класифікації, що робить цей метод надзвичайно потужним для задач з високою розмірністю даних.

2. Штучні нейронні мережі. ШНМ імітують процес навчання мозку за допомогою структур, які складаються з шарів нейронів. Кожен нейрон застосовує нелінійну функцію активації, наприклад, функцію сигмоїду або ReLU:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

або

$$f(x) = \max(0, x) \quad (3)$$

Ці функції допомагають мережі обробляти складні патерни в даних і є основою багатьох сучасних моделей машинного навчання.

3. Нечіткі системи. Нечіткі системи використовують правила типу "якщо-то", які дозволяють формалізувати експертні знання в зрозумілій для користувача формі. Наприклад, правило для діагностики ІХС може бути записане як: "Якщо рівень холестерину високий і артеріальний тиск підвищений, то ймовірність ІХС висока". На відміну від ШНМ та МОВ, нечіткі системи забезпечують більш прозорий процес прийняття рішень.

2.4.2 Порівняння моделей для діагностики ІХС

МОВ, ШНМ та нечіткі системи мають різні характеристики, які роблять їх придатними для задач діагностики ІХС:

- Достовірність: ШНМ часто демонструють високу достовірність у складних завданнях класифікації, але можуть бути схильні до перенавчання, особливо при недостатньо великих наборах даних. МОВ також забезпечують високу достовірність, але їх ефективність зменшується в задачах з великою кількістю класів.

- Прозорість: Нечіткі системи мають перевагу в прозорості, оскільки кожне правило легко інтерпретувати, що є критично важливим у медицині. МОВ та ШНМ часто розглядаються як "чорні ящики", що ускладнює розуміння їхніх рішень без спеціальних методів інтерпретації.

- Складність обчислень: ШНМ і нечіткі системи можуть вимагати значних обчислювальних ресурсів для навчання та налаштування, тоді як МОВ зазвичай працюють швидше з великими наборами даних.

2.4.3 Налаштування параметрів

Для підвищення ефективності кожної моделі використовують оптимізацію параметрів. Зокрема, в процесі крос-валідації визначаються

оптимальні гіперпараметри для кожного методу, щоб мінімізувати похибку класифікації. Наприклад, для МОВ необхідно оптимально налаштувати параметр регуляризації C та параметри ядра, які впливають на здатність моделі адаптуватися до складних даних. У ШНМ налаштовуються параметри, такі як кількість шарів, кількість нейронів у кожному шарі, та коефіцієнт навчання.

2.4.4 Обмеження і способи їх подолання

Попри численні переваги, методи машинного навчання мають певні обмеження, які можуть вплинути на діагностичну здатність системи. Наприклад:

- Схильність до перенавчання (ШНМ): Для зменшення перенавчання застосовують методи регуляризації та збільшення даних, що допомагає покращити узагальнюючу здатність моделі.

- Чорний ящик (МОВ, ШНМ): Недостатня прозорість може обмежувати можливість інтерпретації результатів у медичних застосуваннях. Для розв'язання цієї проблеми використовують методи, такі як LIME (Local Interpretable Model-agnostic Explanations) та SHAP (SHapley Additive exPlanations), які допомагають пояснити рішення моделей.

2.4.5 Практичне використання у діагностиці ІХС

У ряді досліджень методи машинного навчання вже продемонстрували свою ефективність у діагностиці серцево-судинних захворювань. Наприклад, ШНМ використовували для класифікації пацієнтів за рівнем ризику ІХС, досягаючи високих показників достовірності завдяки аналізу електрокардіограм (ЕКГ). МОВ також застосовували для прогнозування серцевих подій на основі аналізу багатьох клінічних параметрів, що дозволяє лікарям отримувати швидкі та точні результати діагностики. Нечіткі системи, у свою чергу, виявилися

корисними для інтерпретації складних комбінацій симптомів, що дозволяє врахувати індивідуальні особливості пацієнта.

2.5 Висновки

У другому розділі роботи було сформовано теоретичну базу та розроблено концептуальні підходи до процесу діагностики ішемічної хвороби серця (ІХС) з використанням сучасних методів машинного навчання. Зокрема, було детально розглянуто моделі машинного навчання, які можуть бути ефективно застосовані для діагностики ІХС. До цих моделей належать метод опорних векторів (МОВ), нейронні мережі та нечіткі системи. Кожна з цих моделей має свої унікальні переваги та недоліки, що дозволяє адаптувати їх до специфічних завдань у медичній сфері, зокрема в аналізі складних і багатовимірних медичних даних.

Метод опорних векторів (МОВ) дозволяє створювати високоточні класифікатори, здатні розрізняти різні стани здоров'я пацієнта на основі множини показників. Однак, через його обмежену інтерпретованість, його застосування може бути обмежене в умовах, коли важлива прозорість процесу діагностики. Нейронні мережі, з іншого боку, можуть аналізувати та обробляти значні обсяги даних з великою кількістю ознак, що підвищує достовірність діагностики. Проте вони також мають низьку прозорість, що може ускладнити їх використання лікарями без спеціальної підготовки. Нечіткі системи, у свою чергу, забезпечують баланс між точністю і прозорістю завдяки можливості використання лінгвістичних змінних та нечітких правил, що робить їх зрозумілішими для медичних фахівців.

Також у розділі було розроблено модель процесу діагностики ІХС, яка включає кілька ключових етапів, починаючи від збору і попередньої обробки даних до побудови моделі та оцінки її достовірності. Зокрема, для покращення достовірності та надійності діагностичних рішень запропоновано стратегію ансамблю класифікаторів, що дозволяє підвищити впевненість у діагнозі за

рахунок об'єднання декількох класифікаційних моделей. Це може особливо допомогти в тих випадках, коли індивідуальні моделі дають неоднозначні або менш достовірні результати.

Запропонована модель процесу діагностики ІХС має значний потенціал для впровадження в системи підтримки прийняття рішень (СППР) у медичній практиці. Вона забезпечує як високу достовірність класифікації, так і необхідний рівень прозорості, що дозволяє лікарю не лише правильно діагностувати наявність ІХС, а й краще зрозуміти причинно-наслідкові зв'язки між клінічними показниками та діагнозом. Це, в свою чергу, може сприяти більш ранньому виявленню патології та підвищенню якості лікування.

3 ІНФОРМАЦІЙНА СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ

3.1 Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Сучасні інформаційні системи відіграють ключову роль в автоматизації процесів медичної діагностики, дозволяючи лікарям швидше і точніше приймати рішення. Особливо це актуально для захворювань, що мають складну симптоматику або вимагають використання великих обсягів даних для аналізу. За допомогою автоматизованих систем підтримки прийняття рішень можна значно знизити ймовірність помилок, мінімізувати людський фактор і забезпечити більш індивідуалізований підхід до лікування пацієнтів.

Такі системи розробляються з метою допомоги лікарям у діагностичному процесі, особливо в складних і невизначених випадках, де необхідно обробити велику кількість даних пацієнта та ознак захворювання. Інформаційні системи підтримки прийняття рішень дозволяють зменшити ймовірність помилкових діагнозів, зокрема завдяки автоматизованій обробці медичної інформації та використанню штучного інтелекту, нечіткої логіки, генетичних алгоритмів та інших сучасних підходів.

Архітектура інформаційної системи визначає, з яких ключових компонентів складається система та яким чином ці компоненти взаємодіють між собою. Для успішної роботи системи кожен етап обробки даних має бути чітко визначений: від отримання вихідної інформації до кінцевого прийняття рішення про наявність або відсутність захворювання. Оскільки інформація, що збирається, є неоднорідною і може походити з різних джерел, важливу роль відіграє її правильна інтеграція та обробка. У рамках цього розділу розглянемо основні компоненти інформаційної системи, починаючи з пристроїв збору даних і закінчуючи програмними модулями для аналізу й прийняття рішення.

Інформаційна система підтримки прийняття рішень для діагностики медичних захворювань складається з ряду важливих компонентів, які

забезпечують її функціональність та ефективність. Перший і ключовий компонент системи — це пристрої, що збирають необхідну для діагностики інформацію. Для визначення ішемічної хвороби серця (ІХС), наприклад, використовуються такі пристрої, як електрокардіографи (ЕКГ), комп'ютерні томографи (КТ) та інші неінвазивні методи дослідження. Дані, отримані з цих пристроїв, включають показники серцевого ритму, артеріального тиску, рівня холестерину та інші фізіологічні параметри пацієнта. Кожен пристрій передає дані до системи у вигляді числових або графічних показників, які далі будуть оброблятися для отримання діагнозу.

Дані, що надходять від медичних пристроїв, збираються і попередньо обробляються модулем збору даних. Цей модуль забезпечує правильний формат передавання даних до центральної частини системи. Його завдання полягає в стандартизації даних, оскільки різні пристрої можуть використовувати різні формати. Крім того, модуль збору даних може виконувати попередню фільтрацію інформації, відокремлюючи непотрібні або застарілі показники, що дозволяє скоротити обсяги даних, які потребують подальшої обробки.

Важливим елементом системи є сервер або хмарне сховище, куди збираються і зберігаються всі дані, отримані від медичних пристроїв. Це може бути внутрішній сервер медичного закладу або віддалене хмарне рішення. Використання централізованого сховища дозволяє зберігати великі обсяги даних, які можуть бути доступні для аналізу в будь-який момент. Дані зберігаються в захищеному вигляді, що гарантує конфіденційність інформації про пацієнтів.

Після збору даних настає етап їх аналізу, на якому використовується модуль відбору ознак. Цей програмний компонент відповідальний за вибір найважливіших параметрів, які можуть суттєво вплинути на кінцевий діагноз. У даній системі можна використовувати алгоритми, такі як Sequential Floating Forward Selection (SFFS), що автоматично визначають ознаки з найменшою похибкою класифікації. Модуль відбору ознак допомагає зменшити складність

обробки, скоротивши кількість ознак до тих, які є найбільш значущими для прийняття рішення. Модуль відбору ознак є одним з критичних елементів системи. Його мета — скоротити кількість параметрів, що аналізуються, і залишити лише ті ознаки, які найбільш ймовірно свідчать про наявність захворювання. Цей модуль забезпечує не лише ефективну роботу системи, зменшуючи час обробки даних, але й підвищує достовірність діагнозу, оскільки у процесі аналізу залишається лише найрелевантніша інформація.

Модуль класифікації та діагностики - це головний обчислювальний модуль, який на основі відібраних ознак приймає рішення про наявність або відсутність захворювання. Він використовує методи машинного навчання або нечіткої логіки для класифікації пацієнтів. Наприклад, штучні нейронні мережі (ШНМ) можуть аналізувати комбінації параметрів пацієнта і робити висновки щодо наявності ішемічної хвороби серця. Крім того, система може інтегрувати стратегію ансамблевих класифікаторів (САК), яка дозволяє враховувати думки декількох різних алгоритмів для покращення достовірності діагнозу, особливо у складних випадках. Метод класифікації в системі базується на поєднанні методів штучної нейронної мережі та нечіткої логіки, що дозволяє досягти високої достовірності при обробці складних медичних даних. Використання стратегії ансамблевих класифікаторів дозволяє комбінувати рішення кількох класифікаторів, підвищуючи надійність діагнозу, особливо у випадках невизначеності. Такий підхід віддзеркалює реальну медичну практику, де часто рішення приймається консиліумом лікарів. Використання методів штучного інтелекту в системах підтримки прийняття рішень дозволяє значно підвищити ефективність медичної діагностики. Машинне навчання дозволяє системі адаптуватися до нових даних, постійно покращуючи свої діагностичні можливості. Зокрема, штучні нейронні мережі можуть аналізувати тисячі медичних випадків та виявляти патерни, які можуть бути невидимі для лікаря. Використання великих наборів даних у поєднанні з AI дозволяє виявляти рідкісні захворювання на ранніх етапах, що підвищує шанси на успішне лікування.

Після аналізу даних і діагностики результати передаються на комп'ютер лікаря або безпосередньо на смартфон пацієнта. Результати можуть бути різними в залежності від достовірності і впевненості в діагнозі. Наприклад:

- Пацієнт має ішемічну хворобу серця.
- Пацієнт не має ішемічної хвороби серця.
- Необхідні додаткові дослідження (якщо ступінь впевненості в рішенні є низьким).

Інформаційна система підтримки прийняття рішення функціонує завдяки тісній взаємодії всіх її складових частин. Кожен елемент, починаючи з пристроїв збору даних і закінчуючи модулем діагностики, має свою функцію. Результат діагностики залежить від правильної взаємодії модулів збору, відбору ознак і класифікації, а також від достовірності алгоритмів машинного навчання, які забезпечують надійність системи.

В кінцевому результаті інформаційна система не лише допомагає лікарям приймати правильні рішення, але й забезпечує більш оперативне обслуговування пацієнтів, зменшуючи час очікування на діагноз і підвищуючи ймовірність раннього виявлення захворювань.

Як показано на рисунку 3.1, модуль збору даних передає стандартизовані показники до центрального сховища для подальшої обробки.

Таким чином, побудова інформаційної системи підтримки прийняття рішень для медичної діагностики вимагає чітко визначеної структури, яка забезпечує інтеграцію різних модулів для збору, обробки та аналізу даних. Кожен елемент системи, починаючи з медичних пристроїв і закінчуючи програмними модулями класифікації, відіграє важливу роль у досягненні достовірності та ефективності діагностики.

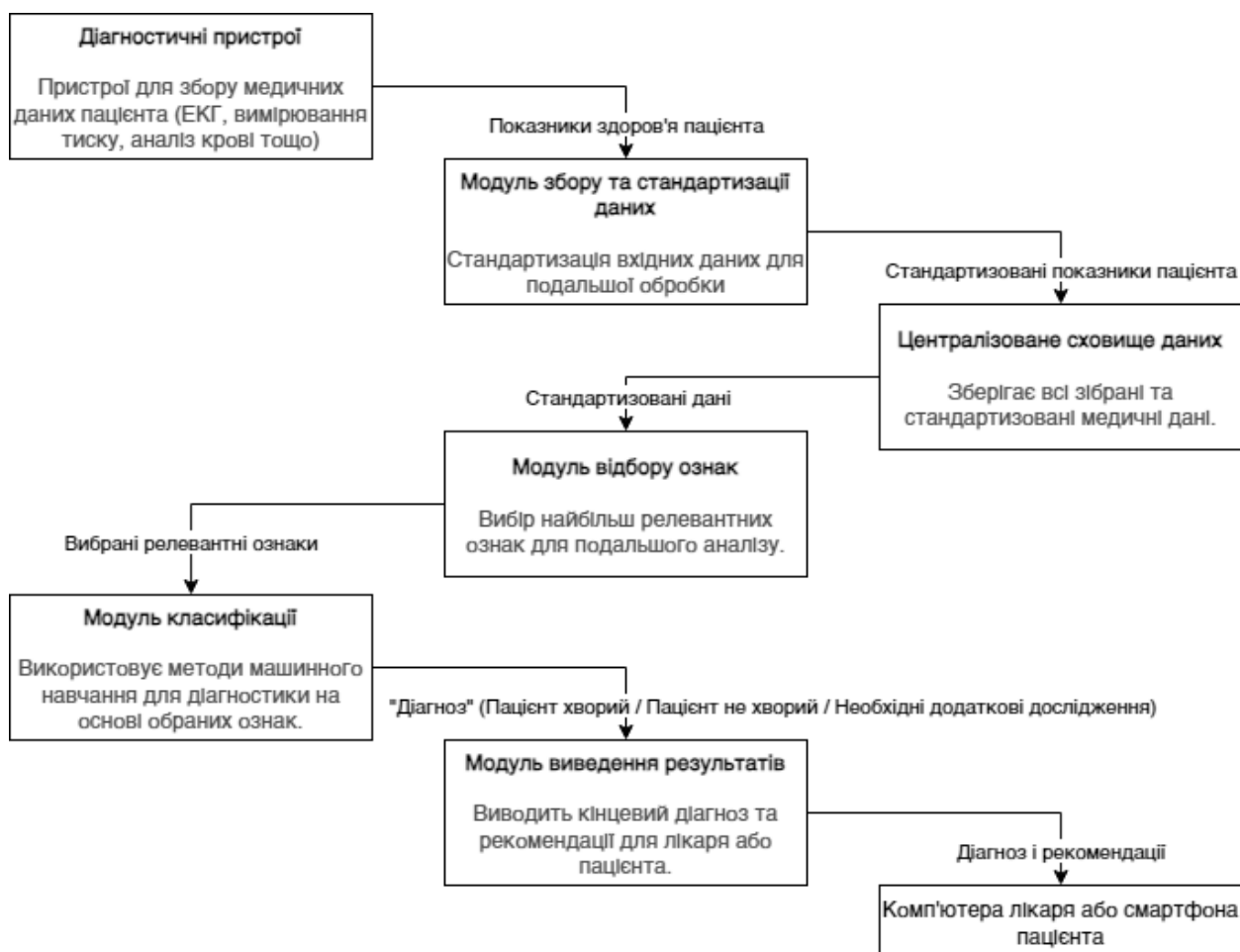


Рисунок 3.1 – Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

У майбутньому, з розвитком технологій штучного інтелекту та великих даних, інформаційні системи підтримки прийняття рішень зможуть не тільки пропонувати діагнози, але й надавати лікарям персоналізовані рекомендації щодо лікування пацієнтів. Інтеграція таких систем з електронними медичними записами та біометричними пристроями може зробити їх ще більш ефективними та доступними для широкого кола медичних закладів.

3.2 Метод визначення медичного діагнозу

Запропонована методологія складається з чотирьох основних етапів:

1. Вибір ознак. Етап вибору ознак передбачає зменшення їх кількості шляхом вибору підмножини, що забезпечує найкращу достовірність класифікації. У цьому дослідженні застосовується метод SFFS, який визначає підмножину з найменшою помилкою класифікації через 10-кратну перехресну перевірку.

2. Генерація системи на основі нечітких правил. Цей етап полягає у розробці класифікатора для медичної діагностики, де лікар визначає наявність ІХС у пацієнта за симптомами. Проблема діагностики формалізується через функцію f , яка відображає набір симптомів a на мітку класу діагнозу c_j .

3. Оптимізація системи на основі нечітких правил. Цей етап полягає у виборі підмножини правил з високою класифікаційною здатністю. Мета оптимізації — максимізувати достовірність f_{acc} та мінімізувати кількість правил f_{rule} за допомогою алгоритму NSGA-II.

4. Оцінка системи на основі нечітких правил. Цей етап включає заходи прозорості, які важко виміряти, але пов'язані з кількістю правил і умов у них, а також використанням лінгвістичних змінних. Для оцінки дискримінаційної здатності системи використовуються два показники: відсоток правильно класифікованих (ВПК) та площа під кривою характеристики роботи приймача (ППК).

Розглянемо детальніше запропоновані кроки детальніше у наступних розділах.

3.2.1 Вибір ознак

Відбір ознак є важливим етапом попередньої обробки даних, що має ключове значення для ефективної роботи інформаційної системи підтримки прийняття рішень. Він полягає у виборі підмножини ознак, що дозволяє

створити найбільш ефективну модель для прогнозування або класифікації, виходячи з певних критеріїв. В умовах великої кількості вхідних даних цей етап допомагає зменшити розмірність задачі, знизити складність обчислень та підвищити продуктивність системи, забезпечуючи при цьому більш точні результати.

Медичні дані, що збираються для діагностики, зазвичай містять велику кількість ознак (параметрів), таких як показники ЕКГ, артеріального тиску, рівня холестерину та інші важливі показники здоров'я пацієнта. Кожна з цих ознак потенційно може мати різний вплив на діагноз, однак не всі з них є рівнозначно важливими. Велика кількість ознак не тільки збільшує час і ресурси, необхідні для обробки даних, але й може призводити до зниження достовірності моделі через явище, яке називається прокляттям розмірності. Це явище виникає тоді, коли кількість параметрів занадто велика, і модель починає враховувати занадто багато несуттєвих або зайвих ознак, що може знизити достовірність класифікації.

Саме тому відбір ознак є критичним етапом попередньої обробки даних. Його метою є зменшення кількості вхідних ознак шляхом вибору лише тих, що мають найбільший вплив на кінцевий результат. Відібрані ознаки забезпечують створення моделі, яка не лише швидко обробляється, але й здатна надати достовірні прогнози або діагнози. Це дозволяє не лише покращити продуктивність системи, але й підвищити її достовірність та надійність.

Серед найпоширеніших методів відбору ознак у науковій та інженерній спільноті можна виділити алгоритм Sequential Forward Floating Selection (SFFS) та генетичні алгоритми (ГА). Обидва ці підходи зарекомендували себе як ефективні інструменти для зменшення кількості ознак та створення найкращих моделей у задачах класифікації і прогнозування.

В даній роботі для відбору ознак використаємо метод SFFS.

Метод SFFS є варіантом класичного методу послідовного пошуку, який дозволяє динамічно додавати та видаляти ознаки в процесі пошуку. На кожному етапі метод додає ознаку, що найбільш покращує модель, а також

може видаляти раніше обрані ознаки, якщо вони виявляються менш корисними на наступних етапах. Це дозволяє методу адаптивно коригувати набір ознак і знаходити найбільш релевантні з них для конкретної задачі. Відмінністю SFFS є його здатність автоматично "плавати" між різними підмножинами ознак, знижуючи ймовірність пропуску важливих ознак або включення зайвих.

Метод SFFS є відносно швидким і менш ресурсоємним у порівнянні з деякими іншими методами відбору ознак. Його алгоритм дозволяє швидко скорочувати простір ознак, роблячи його придатним для задач, де необхідна оперативна обробка великої кількості вхідних даних.

Генетичні алгоритми належать до класу еволюційних методів оптимізації, що імітують процес природного відбору. Вони починаються з випадкових наборів ознак і поступово "еволюціонують", на кожному етапі обираючи найкращі набори на основі критерію, такого як достовірність класифікації. Генетичні алгоритми імітують процеси мутацій і схрещувань, створюючи нові набори ознак, що мають потенціал покращення моделі.

Незважаючи на високу ефективність ГА, цей метод є більш обчислювально складним та ресурсоємним, оскільки потребує значного часу для еволюції наборів ознак до оптимального стану. Однак ГА є потужним інструментом для знаходження глобально оптимальних рішень у складних просторах пошуку, де інші методи можуть застрягти на локальних оптимумах.

Хоча обидва методи є ефективними, у науковій літературі багато досліджень вказують на переваги SFFS над генетичними алгоритмами у задачах класифікації. Основною перевагою SFFS є його швидкість та адаптивність. Метод SFFS може швидко ідентифікувати найбільш релевантні ознаки, що робить його придатним для задач, де час обробки є критичним фактором. З іншого боку, генетичні алгоритми краще підходять для задач, де потрібно шукати глобально оптимальне рішення, але вони вимагають більше обчислювальних ресурсів.

Досі немає єдиного висновку, який з цих двох методів є найкращим у кожній конкретній задачі. Кожен з них має свої переваги і недоліки, залежно

від контексту використання. У цьому дослідженні обрано метод SFFS, оскільки він дозволяє швидше та ефективніше скоротити набір ознак, що є важливим для оперативного аналізу медичних даних.

Стандартний алгоритм Sequential Forward Search (SFS) розпочинається з порожнього набору ознак і поступово формує підмножини ознак-кандидатів за наступною процедурою:

1. Виконується оцінка всіх наборів ознак розміру 1, після чого обирається ознака з найвищою ефективністю.
2. Оцінюються всі можливі набори ознак розміру 2, які включають раніше обрану ознаку, та обирається оптимальна комбінація.
3. Оцінюються набори ознак розміру 3, що містять ознаки з попереднього кроку, і знову обирається найбільш ефективна комбінація.
4. Процес триває за аналогічною схемою, доки не буде досягнуто встановленого критерію зупинки.

Кількість обраних ознак може бути визначена як користувачем, так і автоматично налаштована самим алгоритмом.

Метод Sequential Floating Forward Selection (SFFS) вдосконалює SFS завдяки можливості динамічного коригування кількості ознак, які можуть бути додані або видалені на кожному кроці. Це дозволяє алгоритму переглядати рішення на попередніх етапах, що дає можливість покращувати результати вибору ознак.

Sequential Forward Selection (SFS) і Sequential Backward Selection (SBS) є основними методами вибору ознак у машинному навчанні, які спрямовані на покращення продуктивності моделей. Вони оптимізують набір ознак, поступово додаючи або видаляючи їх. На рисунку 3.2 зображено алгоритм SFFS.

SFS починає з порожнього набору ознак і на кожному етапі додає ту ознаку, яка найбільше покращує модель. Це протилежність SBS, який починає з повного набору ознак і послідовно видаляє найменш значущі для моделі ознаки, доки не буде досягнуто оптимального результату. Обидва методи

корисні для зменшення розмірності простору ознак, що часто призводить до більш ефективних і достовірних моделей.

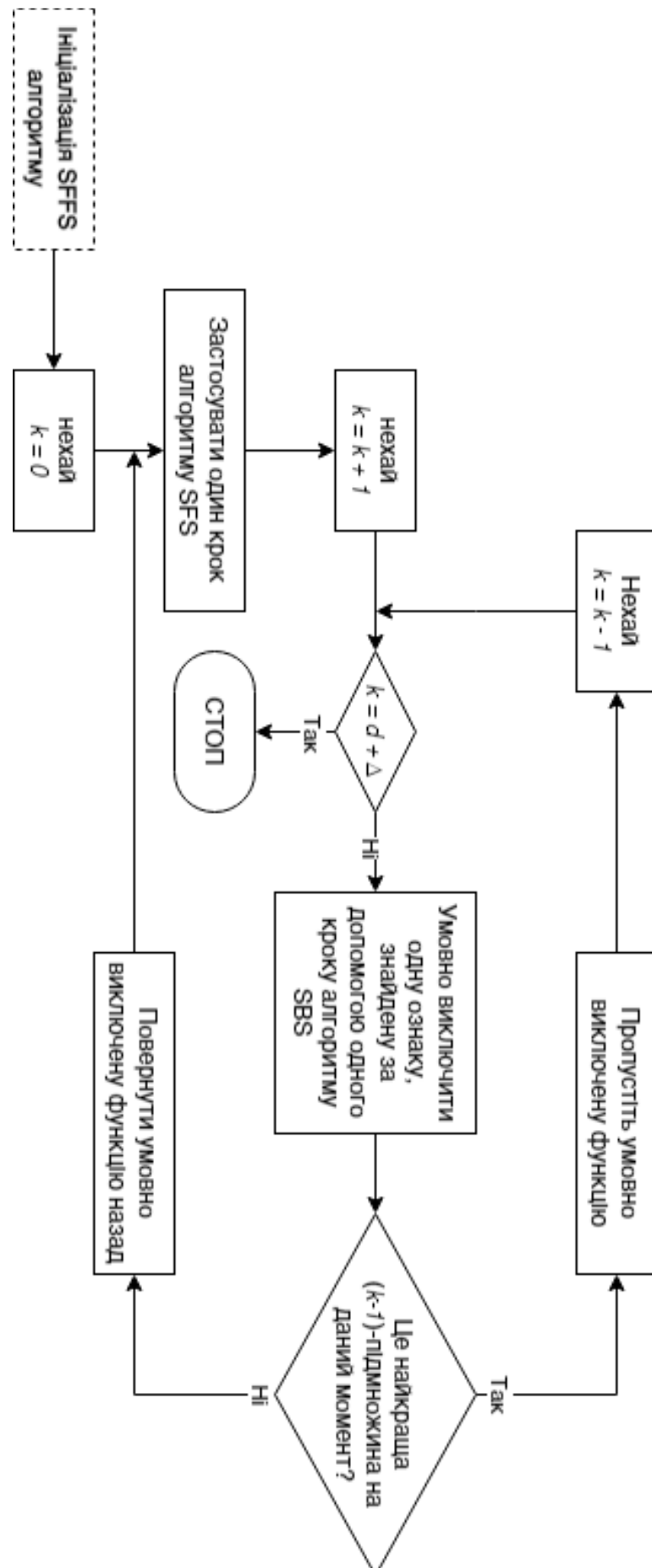


Рисунок 3.2 – Процес Sequential Floating Forward Selection (SFFS)

Принцип роботи алгоритму Sequential Forward Floating Selection (SFFS) представлений у вигляді наступної схеми:

1. Початковий набір ознак — алгоритм починає з порожнього набору ознак і поступово додає ознаки, які найбільше підвищують достовірність моделі.

2. Оцінка наборів ознак — на кожному кроці додаються або видаляються ознаки на основі того, як вони впливають на критерій достовірності класифікації. Якщо додавання нової ознаки значно покращує модель, вона залишається у наборі. Якщо ж нова ознака не покращує або навіть погіршує модель, алгоритм видаляє її і повертається до попередніх кроків для переоцінки інших ознак.

3. 10-кратна перехресна перевірка — кожна підмножина ознак перевіряється за допомогою 10-кратної перехресної перевірки для оцінки її достовірності. Це дозволяє оцінити, наскільки добре вибрані ознаки дозволяють моделі класифікувати нові дані, зменшуючи ймовірність перенавчання.

4. Критерій зупинки — процес триває доти, поки не буде досягнуто оптимальної кількості ознак або не буде більше ознак, що покращують модель.

5. Цей підхід дозволяє знайти підмножину ознак, яка забезпечує найкращу достовірність класифікації при мінімальній кількості параметрів, що робить модель швидшою і достовірнішою.

Важливо зазначити, що одним із ключових аспектів вибору ознак є критерій, на основі якого приймається рішення про релевантність тієї чи іншої ознаки. У задачах класифікації найчастіше використовується достовірність класифікації як основний критерій. Достовірність класифікації визначається як співвідношення правильно класифікованих випадків до загальної кількості випадків. Чим вища достовірність, тим краще модель здатна відрізнити різні класи захворювань на основі наявних ознак.

Кожна підмножина ознак-кандидатів перевіряється за допомогою перехресної перевірки, і алгоритм SFFS обирає підмножину, яка дає найменшу помилку класифікації. Це означає, що обрані ознаки є найбільш

інформативними для визначення діагнозу і допомагають зменшити ризик неправильних діагнозів.

Основною перевагою SFFS є його здатність динамічно коригувати набір ознак та швидко знаходити оптимальну підмножину ознак, що робить цей метод ефективним для великомасштабних задач. Крім того, завдяки механізму "плаваючого" вибору ознак, він дозволяє уникнути ситуацій, коли важливі ознаки були видалені на ранніх етапах пошуку.

Однак SFFS також має деякі обмеження. Наприклад, він може бути менш ефективним у задачах з високою розмірністю простору ознак, де кількість можливих комбінацій ознак є надто великою для ефективного пошуку. У таких випадках генетичні алгоритми можуть виявитися більш підходящими, оскільки вони здатні знаходити глобальні оптимальні рішення у великих просторах.

3.2.2 Модель процесу генерації нечітких правил

У задачі медичної діагностики лікар на основі симптомів, що спостерігаються у пацієнта, відносить хворобу пацієнта до одного з двох класів, а саме: пацієнт з ІХС та пацієнт без ІХС (бінарна класифікація). Таким чином, проблема медичного діагнозу відноситься до проблем класифікації.

Проблему діагностики можна вирішити, знайшовши відповідний класифікатор або математичну функцію f , яка відображає набір симптомів a на мітку діагностичного класу c_j . Ця функція може бути записана наступним чином:

$$f: a \rightarrow c_j \quad (4)$$

де $a = \{a_1, a_2, \dots, a_n\}$ - множина ознак, а c_j де $j \in \{1, 2\}$ - мітка класу. Класифікатором може бути ШНМ, дерево рішень або будь-який інший метод класифікації.

У теорії нечітких множин класифікатор f - це набір нечітких правил, де k -те правило має наступний формат:

$$R_k : \text{якщо } a_1 \in B_1^k \text{ та } a_2 \in B_2^k \text{ та } \dots \text{ та } a_n \in B_n^k \text{ тоді } Y \in c_j \quad (5)$$

де B_n^k - нечіткі множини вхідних змінних (a_1, a_2, a_n) , представлені лінгвістичними значеннями, такими як низький, середній та високий, тоді як c_j - мітка класу змінної класу Y (яка в цьому випадку є або $Y = c_1$ для наявного класу, або $Y = c_2$ для відсутнього класу).

У деяких дослідженнях використовується розширений формат (5), який включає ступінь визначеності або впевненості r_k k -го правила, де $r_k \in [0, 1]$. Значення r_k представляє також ступінь достовірності рішення, прийнятого за цим правилом. Ця особливість дозволяє лікарю знати про ступінь достовірності рішення, прийнятого нечітким класифікатором, а точніше правилом-переможцем, і чи варто зважати на нього чи ні. Цей формат системи на основі нечітких правил можна записати наступним чином:

$$R_k : \text{якщо } a_1 \in B_1^k \text{ та } a_2 \in B_2^k \text{ та } \dots \text{ та } a_n \in B_n^k \text{ тоді } Y \in c_j \text{ з } \quad (6)$$

достовірністю r_k

де r_k - ступінь достовірності k -го правила.

Хоча (6) дає більше інформації про k -те правило, ніж (5), вона не дозволяє особі, яка приймає рішення, знати ступінь важливості правила і точно, скільки разів правило R_k справджується в навчальних даних. Ця концепція відома в галузі інтелектуального аналізу даних як підтримка і є важливим критерієм для оцінки правил. Використання лише ступеня достовірності може ввести в оману при прийнятті рішення, наприклад, припустимо, що ступінь достовірності R_k дорівнює $r_k = 1$ (100% достовірність класифікації), але він охоплює лише дуже мало випадків у навчальних даних. Як наслідок, R_k має низьку підтримку і не може використовуватися самостійно для узагальнення зв'язків або висновків.

Крім того, таке значення підтримки допомагає впоратися з частковими протиріччями між правилами та розглянути особливі випадки. У цьому дослідженні, для того, щоб надати більше інформації про правила, формат нечітких правил розширено за рахунок включення значення підтримки правила. Цей розширений формат можна записати наступним чином:

$$R_k : \text{якщо } a_1 \in B_1^k \text{ та } a_2 \in B_2^k \text{ та } \dots \text{ та } a_n \in B_n^k \text{ тоді } Y \in c_j \text{ з } \quad (7)$$

достовірністю r_k та підтримкою s_k

де s_k - значення підтримки k -го правила.

3.2.3 Процедура генерації нечітких правил

Процедура генерації нечітких правил необхідна для формування бази знань, що використовуватиметься в процесі прийняття рішення в системі. Нечіткі правила дозволяють системі оперувати неточними або невизначеними даними, що часто трапляється в медичних діагнозах, і приймати рішення на основі нечіткої логіки. Це дає змогу більш гнучко і точно обробляти складні медичні дані, забезпечуючи більш достовірні результати діагностики навіть за відсутності чітко визначених меж між станами пацієнта (хворий/здоровий).

Нехай L_i - кількість лінгвістичних значень, пов'язаних з вхідною змінною a_i . Одне з цих лінгвістичних значень використовується як нечіткий антецедент B_i^k для вхідної змінної a_i в кожному правилі R_k . Крім того, «неважливе» є додатковим лінгвістичним значенням, яке представляє нерелевантні нечіткі антецеденти, які можуть бути видалені без впливу на продуктивність нечіткої системи. Таким чином, кількість N можливих комбінацій антецедентної частини в цьому випадку становить:

$$N = (L_1 + 1) * (L_2 + 1) * (L_3 + 1) * \dots * (L_n + 1) \quad (8)$$

За допомогою цього методу можна згенерувати всі можливі антецеденти нечітких правил, але коли кількість вхідних даних велика (набір даних високої розмірності), кількість згенерованих правил буде експоненціально збільшуватися. Оскільки застосовується процедура виділення ознак на етапі попередньої обробки, кількість вхідних даних буде достатньою для застосування цього підходу. Після визначення антецедентів нечітких правил можна обчислити клас правила R_k , його ступінь визначеності та підтримку наступним чином:

1. Обчислення відповідного класу.

Для кожного навчального прикладу $a_t = (a_{t1}, a_{t2}, \dots, a_{tn})$ обчислюється оцінка його сумісності $\mu_k(a_t)$ з правилом R_k як:

$$\mu_{R_k}(a_t) = \mu_{k1}(a_{t1}) * \mu_{k2}(a_{t2}) * \dots * \mu_{kn}(a_{tn}), \quad t = 1, \dots, m \quad (9)$$

де $\mu_{ki}(\cdot)$ - функція належності нечіткої множини χ_j^k .

Для кожного класу обчислити суму оцінок сумісності навчальних прикладів з нечітким правилом R_k як:

$$\chi_{Class\ h}(R_k) = \sum_{x_t \in Class\ h} \mu_{R_k}(a_t), \quad h = 1, 2 \quad (10)$$

Знайти клас c_j , який має максимальне значення $\chi_{Class\ h}(R_k)$:

$$\chi_{Class\ c_j} = \max\{\chi_{Class\ 1}(R_k), \chi_{Class\ 2}(R_k)\} \quad (11)$$

Якщо клас, що впливає з нечіткого правила, не може бути однозначно визначений, нечітке правило R_k не генеруємо.

2. Обчислення ступеня достовірності

Ступінь достовірності можна розрахувати наступним чином:

$$CF_j = \left\{ \chi_{class\ c_j}(R_k) - \bar{\chi} \right\} / \sum_{h=2}^2 \chi_{class\ h}(R_k) \quad (12)$$

де

$$\bar{\chi} = \sum_{h=1, h \neq c_k}^2 \chi_{class\ h}(R_k) \quad (13)$$

3. Обчислення підтримки

Обчислення підтримки s_k за наступною формулою:

$$s_k = \frac{NCC_{R_k}}{m_j}, R_k \in FRBS_{c_j} \quad (14)$$

Де NCC_{R_k} - кількість навчальних прикладів, правильно класифікованих за правилом R_k , m_j - кількість навчальних прикладів, що належать до класу c_j , а $FRBS_{c_j}$ - множина нечітких правил, пов'язаних з класом c_j .

Нечітке міркування, прийняте в цьому дослідженні, базується на єдиному переможці правила. Новий шаблон $a_t = (a_{t1}, a_{t2}, \dots, a_{tn})$ віднесено до наступного класу переможця R_ω . Переможцем є правило, яке має найбільшу комбінацію між ступенем збігу шаблону з попередньою частиною та ступенем визначеності класів. Правило-переможець можна визначити наступним чином:

$$\mu_{R_\omega}(a_t).CF_\omega = \max \left\{ \mu_{R_j}(a_t).CF_k \mid R_k \in FRBS \right\} \quad (15)$$

Класифікація нового шаблону відхиляється у випадку, коли два або більше правил з різними класами мають однакове максимальне значення в рівнянні 11.

3.2.4 Оптимізація системи на основі нечітких правил

Після генерації всіх можливих правил, процедура оптимізації застосовується для пошуку підмножини з невеликої кількості правил (в межах всіх можливих правил) з високою класифікаційною здатністю. Перша мета представляє достовірність НСНП, тоді як прозорість (підмножина з невеликою кількістю правил) визначає другу мету. Ці дві цілі моделювання нашої НСНП можна записати наступним чином:

$$\text{Maximize } f_{acc}(\text{НСНП}), \text{Minimize } f_{rule}(\text{НСНП}) \quad (16)$$

Де $f_{acc}(\text{НСНП})$ – достовірність НСНП, що вимірюється показником достовірності класифікації, $f_{rule}(\text{НСНП})$ - кількість нечітких правил НСНП. Для одночасної оптимізації вищезазначених цілей використовується керований NSGA-II, а результатами оптимізації є Парето-фронт розв'язків, які представляють ряд різних розв'язків НСНП відповідним їм 2-кортежем значень t_{FRBS} , де $t_{\text{FRBS}} = (\text{достовірність НСНП}, \text{кількість правил у НСНП})$. Вибір між цими розв'язками зазвичай залежить від уподобань користувача та типу досліджуваної задачі. Оскільки достовірність у медичній діагностиці є критичною, було обрано НСНП найвищою достовірністю. У наступному підрозділі наведено короткий опис використаного багатоцільового генетичного алгоритму.

Генетичні алгоритми - це евристичні методи, натхненні природною еволюцією для пошуку оптимального рішення. Багатоцільові генетичні алгоритми (БЦГА) - це клас генетичних алгоритмів, які в основному застосовуються для оптимізаційних задач, що мають декілька і навіть

суперечливі цілі. БЦГА використовують два підходи до розв'язання багатоцільових задач оптимізації. У першому підході різні цільові функції об'єднуються в одну функцію за допомогою вагових коефіцієнтів. Складність цього підходу полягає у визначенні правильних вагових коефіцієнтів, які характеризують вподобання користувача. Другий підхід шукає не домінуючі оптимальні за Парето компроміси між конфліктуючими цілями. Він є більш практичним, оскільки пропонує особі, яка приймає рішення, кілька альтернатив, які мають однакову вартість.

NSGA-II - це ефективний MOGA, запропонований Дебом та ін. для подолання деяких недоліків NSGA, таких як складність обчислень, необхідність вказівки параметра розподілу та підхід, що не передбачає елітарності. Перевагами NSGA-II порівняно з іншими MOGA є збереження різноманітності та швидке недомінантне сортування особин. Поняття недомінантного відношення можна визначити наступним чином. Розв'язок S_A домінує над S_B , якщо виконуються наступні дві умови:

1. S_A строго краща за S_B принаймні в одній цілі
2. S_A не гірша за S_B в усіх цілях

Це вдосконалена версія NSGA-II для контролю ступеня елітарності до певної частки, визначеної користувачем. Цей підхід забезпечує кращу збіжність порівняно з оригінальним NSGA-II. У наступному підрозділі описано представлення хромосом, що використовується у нашій реалізації.

Нехай N - загальна кількість згенерованих нечітких правил, а S - підмножина обраних правил, де $S \leq N$. Хромосома кодується як двійковий рядок довжиною N , і кожен двійковий біт представляє одне правило. Коли значення q -го біта дорівнює 1, вибирається q -те правило, а в іншому випадку (q -й біт дорівнює 0), q -те правило не розглядається. На рисунку 1 показано приклад кодування хромосоми для процедури вибору правил із загальною кількістю правил $N = 10$ та кількістю вибраних правил $S = 5$. Правила під номерами 1,5,6,7 та 10 обрано для формування системи-кандидата на основі

нечітких правил, тоді як правила під номерами 2,3,4,8 та 9 проігноровано (рисунок 3.3).

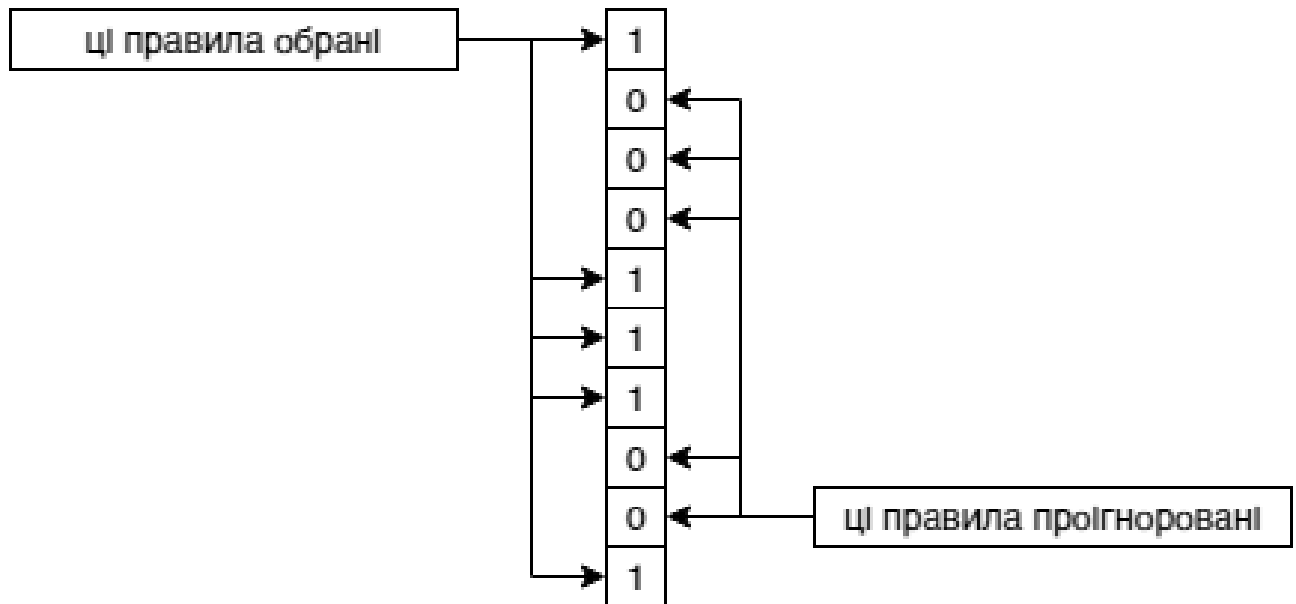


Рисунок 3.3 – Хромосомне кодування процедури вибору правил з $N = 10$, $S = 5$

Підвищення достовірності класифікації системи на основі нечітких правил за допомогою стратегії ансамблевих класифікаторів (САК)

Після створення системи на основі нечітких правил запропоновано підхід до підвищення продуктивності НСНП шляхом підтримки її рішення (класифікації) ансамблем класифікаторів у випадку, коли рішення про класифікацію приймається за правилом (правилом-переможцем), ступінь достовірності якого є нижчим за поріг α , де α визначається користувачем, у цьому дослідженні $\alpha = 0,5$.

Метод Ensemble використовує набір відносно достовірних і різноманітних класифікаторів для класифікації всіх нових випадків шляхом зваженого голосування їх класифікацій. У цьому випадку використовуємо для класифікації НСНП і не викликаємо метод Ensemble у всіх випадках (як це робиться традиційно), а лише у випадках невизначеності, тобто там, де впевненість класифікації НСНП є низькою. Причина, по якій пропонується цей метод, може бути сформульована наступним чином: у випадках, коли впевненість класифікації НСНП є високою (класифікація НСНП є найбільш

вірогідною), вважається за краще продовжувати використовувати НСНП для отримання інтерпретації випадків, в той час як у випадках невизначеності (ступінь впевненості НСНП є низькою) використовується метод Ensemble (метод мішків з простою більшістю голосів), щоб підвищити вірогідність їх правильної класифікації.

Для складання ансамблю класифікаторів було обрано, на додаток до НСНП, шість інших популярних класифікаторів, з яких 3 є нечіткими методами, а решта - нечіткими методами. До нечітких методів відносяться Штучні нейронні мережі (ШНМ), популярний алгоритм дерева рішень C4.5 та статистичний метод лінійного дискримінантного аналізу (ЛДА), а до нечітких методів належать FH-GBML, SLAVE та GP-FCS. Ці алгоритми включені до багатьох некомерційних інструментів, таких як Weka, Keel та Orange.

3.2.5 Оцінка системи на основі нечітких правил

Хоча концепцію прозорості важко перетворити на вимірювані метрики, багато дослідників погоджуються з деякими властивостями прозорості, пов'язаними з нечіткими системами, такими як кількість правил та кількість попередніх умов у кожному правилі. Крім того, особливий акцент робиться на використанні лінгвістичних змінних, які мають чітке значення. У цьому дослідженні розглядаються всі вищезазначені заходи, а також підкреслюємо важливість надання більшої кількості інформації про ступінь визначеності та підтримки кожного правила, що дозволить створювати прозорі та інформативні правила.

Для оцінки дискримінаційної здатності системи на основі нечітких правил використовуються два добре відомі методи. Першим і традиційним методом є відсоток правильно класифікованих (ВПК) тестових патернів, а другим - площа під кривою характеристики роботи приймача (ROC) або скорочено ППК. Останній показник широко використовується в медичній спільноті для оцінки діагностичної потужності тестів на захворювання. Крім

того, для оцінки достовірності класифікації нашої системи на основі нечітких правил використовується метод 10-кратної перехресної перевірки. Результати, досягнуті цією системою, порівнюються з існуючими еталонними методами.

3.3 Висновки

У третьому розділі було детально розглянуто структуру та основні компоненти інформаційної системи підтримки прийняття рішення для діагностики медичних захворювань. Ключову роль у роботі такої системи відіграють як апаратні компоненти, що забезпечують збір даних, так і програмні модулі, які виконують їх обробку та аналіз. Першочерговим завданням інформаційної системи є збір і стандартизація медичних даних пацієнтів, отриманих від різних діагностичних пристроїв, таких як ЕКГ, КТ, ПЕТ та інші. Після цього дані передаються до модулів, які займаються відбором ознак та класифікацією.

Відбір ознак є важливим етапом попередньої обробки даних, що дозволяє скоротити кількість параметрів, з якими працює система, і зосередитися на найбільш релевантних для діагностики. У нашій системі було використано метод Sequential Forward Floating Selection (SFFS), який, завдяки своїй швидкості та адаптивності, дозволяє ефективно обрати підмножину ознак, що підвищує достовірність класифікації.

Модуль класифікації, що є основним обчислювальним блоком системи, використовує сучасні методи машинного навчання, такі як штучні нейронні мережі та нечітка логіка, для прийняття рішення про наявність або відсутність захворювання. Цей модуль є вирішальним, оскільки саме він генерує кінцевий діагноз, на основі якого лікар може приймати рішення щодо подальших дій.

Таким чином, структура інформаційної системи підтримки прийняття рішення побудована таким чином, щоб забезпечити високу достовірність діагностики, швидкість обробки даних і можливість адаптації під різні медичні задачі. Описані модулі та їх взаємодія створюють надійний інструмент для

допомоги лікарям у процесі прийняття рішень, підвищуючи ефективність медичних обстежень і знижуючи ймовірність помилок.

У майбутньому можливе вдосконалення інформаційної системи шляхом інтеграції з більш сучасними технологіями обробки даних та машинного навчання. Наприклад, використання глибоких нейронних мереж або систем на основі великих даних може суттєво підвищити достовірність діагностики, особливо у випадках рідкісних захворювань. Окрім цього, система може бути розширена для автоматичного аналізу зображень (наприклад, з томографів чи рентгенівських знімків), що дозволить зменшити навантаження на лікарів та пришвидшити процес обробки інформації. Інтеграція з портативними медичними пристроями, які збирають дані в режимі реального часу (пульс, тиск, рівень кисню в крові), зробить систему ще більш мобільною та доступною для пацієнтів.

4 ОЦІНКА ЕФЕКТИВНОСТІ ІНФОРМАЦІЙНОЇ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ

4.1 Постановка експериментальних досліджень

У цьому розділі буде представлено опис створення та використання прототипу інформаційної системи, розробленого для підтримки прийняття рішень у медичній діагностиці. Основним завданням цієї системи є надання допомоги медичним фахівцям у процесі діагностування захворювань на основі великої кількості медичних даних, що збираються та обробляються автоматизованими засобами. Прототип реалізовано у вигляді програмного забезпечення, яке поєднує в собі сучасні алгоритми обробки даних та класифікації для підвищення достовірності діагностики та скорочення часу на прийняття рішень.

Важливо зазначити, що розробка прототипу інформаційної системи не лише відображає теоретичні знання у галузі обробки даних, але й представляє практичну реалізацію цих знань у вигляді програмного продукту. Ця система дозволяє лікарям працювати з даними більш ефективно, мінімізуючи людський фактор та зменшуючи ймовірність помилок при аналізі складних медичних випадків. Програмне забезпечення також включає модулі для автоматизованого вибору релевантних ознак, що суттєво впливають на діагностику.

Для оцінки ефективності розробленого прототипу буде проведено серію експериментальних досліджень. Експерименти передбачають використання системи для аналізу медичних даних пацієнтів, що будуть взяті з різних джерел та охоплюватимуть широкий спектр захворювань. Кожен експеримент буде спрямований на оцінку того, як прототип інформаційної системи працює в різних умовах і як різні алгоритми обробки та класифікації впливають на достовірність і швидкість діагностики.

Для цього будуть використані різні набори даних та алгоритми машинного навчання, зокрема ті, що були обрані на основі теоретичного

аналізу у попередніх розділах роботи. В експериментах будуть порівнюватися показники достовірності класифікації, продуктивність системи, швидкість обробки даних та її стійкість до змін у вхідних даних. Окремо буде проведено оцінку стабільності роботи програмного забезпечення при роботі з великими обсягами медичної інформації.

Система підтримки прийняття рішень у медичній діагностиці, реалізована у вигляді прототипу програмного забезпечення, також дозволяє легко адаптуватися до різних типів захворювань і медичних даних, завдяки модульній структурі і можливості масштабування. Це відкриває перспективи подальшого вдосконалення прототипу і його впровадження в реальні медичні установи для допомоги лікарям у прийнятті складних діагностичних рішень. Експерименти, проведені на базі цієї системи, дозволять чітко оцінити її ефективність і визначити напрямки для подальшого розвитку та оптимізації.

Таким чином, результати проведених експериментів допоможуть сформулювати висновки щодо можливих покращень прототипу, зокрема в аспектах підвищення достовірності діагностування, зменшення часу на обробку даних та підвищення загальної стабільності системи. Використання різних підходів до обробки даних і порівняння їх результатів дозволить отримати комплексну картину можливостей та обмежень прототипу, що стане основою для наступних етапів розробки та вдосконалення інформаційної системи.

4.2 Процедура вибору ознак

На цьому кроці метод SFFS було застосовано для вибору підмножини ознак, які можуть створити найкращу модель класифікації. Результати цього кроку наведено в таблиці 4.1. Як видно з таблиці, SFFS обирає підмножину, яка включає наступні чотири ознаки: стать, тип болю в грудях, кількість магістральних судин і таласемію. Крім того, хоча кількість ознак зменшується з 13 до 4, достовірність класифікації моделі зберігається на рівні майже 83%, що свідчить про здатність методу SFFS справлятися з проблемою вибору ознак.

Таблиця 4.1 – Результати процедури відбору ознак

	Всі можливості	Вибрані можливості
Кількість можливостей	13	4
Опис	Див. таблицю 1.1	2 (стать), 3 (тип болю в грудях), 12 (кількість магістральних судин), 13 (таласемія)
Помилка класифікації	17.07%	17.04%

4.3 Нечітка оптимізація

Цей крок має на меті:

1. Максимізувати достовірність класифікації
2. Мінімізувати кількість нечітких правил у системі на основі нечітких правил.

Результати цього кроку наведено в таблицях 4.2, 4.3, В.1, В.2, В.3 та В.4.

Позначення S1...S10 у таблиці 4.2 представляють підмножини ознак, що використовуються для побудови моделі класифікації та генерування нечітких правил. Це підмножини всіх доступних ознак, які були відібрані на попередньому етапі обробки даних, зокрема через процедуру відбору ознак.

Таблиця 4.2 – кількість нечітких правил та нечітких множин, обраних на кроці оптимізації

Підмножини	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Середнє значення
кількість правил	7	8	8	10	8	10	6	7	7	9	8
кількість початкових нечітких множин	28	32	32	40	32	40	24	28	28	36	32
кількість обраних нечітких множин	17	22	20	27	20	25	15	15	17	21	19.9
кількість нечітких правил	2.4	2.8	2.5	2.7	2.5	2.5	2.5	2.1	2.4	2.5	2.5

4.3.1 Критерій прозорості

Загальна кількість згенерованих нечітких правил становить $N = 300$ правил, з яких для представлення нечіткої системи на основі правил відбирається S правил. Результатом цього відбору є розв'язки, що не домінують за Парето, які представляють нечіткі системи на основі правил з різними рівнями достовірності кількості значень правил. Як зазначалося раніше, пріоритет надається достовірності систем, тобто обирається НСНП з вищою достовірністю. У таблиці 4.2 наведено кількість правил, нечітких множин та кількість множин на одне правило для різних згинів. Як видно з таблиці 4.2, нечіткі правила, вибрані з 300 правил, мають відносно невелику кількість - в середньому 8 правил на одне згини. Різниця в кількості правил між складками пояснюється зміною навчальних даних від складки до складки. Ще одним

покращенням є кількість нечітких множин, що обираються для кожного правила, яка зменшилася з 4 до 2,5 на правило. Крім того, нечіткі правила використовують зрозумілий опис попередніх умов у вигляді правил типу «якщо-то» і включають в себе, крім міток класу, ступінь їх визначеності та підтримки. Лінгвістичний опис допомагає лікарю зрозуміти зв'язок між факторами або симптомами і результатом діагнозу, в той час як ступінь достовірності допомагає зрозуміти, наскільки цей зв'язок є точним. Цей тип міркувань відповідає людському мисленню, яке допускає існування часткової істини.

З іншого боку, підтримка правила дозволяє дізнатися, до якої міри це відношення можна узагальнити. Наприклад, у таблиці 4.3, 8-е правило має ступінь достовірності 1, що означає ідеальну здатність до класифікації, хоча воно охоплює лише 2 випадки в навчальних даних. Такого роду зв'язки можна розглядати як окремі випадки і не можна їх узагальнювати. З іншого боку, правила 1, 2 і 3 правильно класифікували 82, 56 і 30 випадків відповідно з 243 випадків (загальна кількість випадків у навчальних даних). Ці правила повинні мати особливе значення при аналізі та формулюванні висновків у порівнянні з іншими, оскільки вони займають домінуючу позицію в процесі діагностики.

4.3.2 Критерій достовірності класифікації

Таблиці В.1 і В.2 показують відповідно значення ВПК і ППК для нашої НСНП до застосування стратегії ансамблевих класифікаторів (САК), яку в таблиці названо Д-НСНП, і після застосування цієї стратегії (П-НСНП). Крім того, в таблицях В.1 і В.2 наведено також значення ВПК і ППК для ряду еталонних методів, а саме: (1) дерево рішень C4.5, (2) багат шаровий перцептрон (MLP), який є найпопулярнішою архітектурою керованих нейронних мереж, (3) HF-GBML, нечіткий гібридний генетичний метод машинного навчання для генерації інтерпретованих нечітких правил, (4) SLAVE, метод навчання на основі генетичного алгоритму для генерації

нечітких правил, (5) GP FCS, алгоритм генетичного програмування для оптимізації нечітких класифікаторів, та (6) статистичний метод лінійного дискримінантного аналізу (ЛДА).

Таблиця 4.3 – список правил класифікації підмножини 2 та відповідні їм значення підтримки, достовірності та достовірності класифікації

Правило	Мітка класу	Достовірність	Підтримка	Тренування			Тестування		
				Дані	Вірно	Хибно	Дані	Вірно	Хибно
1	0	0.28	0.337	101	82	19	9	9	0
2	1	0.80	0.231	62	56	6	9	8	1
3	0	0.72	0.124	35	30	5	3	3	0
4	1	0.27	0.074	20	18	2	3	3	0
5	1	0.74	0.037	12	9	3	1	1	0
6	0	1.0	0.033	8	8	0	2	2	0
7	0	1.0	0.012	3	3	0	-	-	-
8	1	0.20	0.004	2	1	1	-	-	-
				243	207	36	25	26	1

Як видно з таблиць В.1 і В.2, початкові результати показують, що стратегія ансамблевих класифікаторів, застосована для підтримки класифікації правил зі значенням достовірності менше 0.5, покращила достовірність з 81.85% до 84.44% і з 0.812 до 0.839 для ВПК і ППК відповідно. Ця достовірність є порівнянною або навіть кращою, ніж у інших класифікаторів, наведених у таблицях В.1 і В.2. Цей результат також вказує на те, що більш надійне класифікаційне рішення може бути отримане в поєднанні з іншими класифікаторами, відомими своєю достовірністю. Цей висновок поділяють Вест та ін., які припустили, що достовірність набору класифікаторів, як правило, краща, ніж найкращого окремого класифікатора. Насправді, в медичній

практиці поширеною є ситуація, коли є більше однієї думки у випадках, коли ступінь визначеності діагнозу є слабким. Таким чином, ці класифікатори відіграють роль експертів, які надають свої висновки щодо складних випадків, які мають низький ступінь визначеності.

4.3.3 Чутливість і специфічність

ROC-криві НСНП до і після застосування стратегії ансамблевих класифікаторів (САК) показані на рисунку 4.1 синім і зеленим кольорами відповідно. Ці криві є графічним відображенням залежності чутливості від значень (1-специфічності) двох систем. Вони також представляють значення істинно-позитивних (ІП), хибно-позитивних (ХП), істинно-негативних (ІН) та хибно-негативних (ХН) результатів для кожної системи. У нашому випадку позитивним класом є наявність ІХС, а негативним - відсутність ІХС. З таблиць В.3 і В.4, а також рисунку 4.1 видно, що значення специфічності для НСНП явно перевищує значення чутливості. Це спостереження також справедливе для всіх класифікаторів (за винятком алгоритму MLP) і особливо для нечітких систем, таких як SLAVE, HF-GBML і Д-НСНП, де різниця між специфічністю і чутливістю дуже значна, наприклад, специфічність у Д-НСНП становить 0.873, тоді як чутливість – 0.750. Іншим спостереженням є те, що нечіткі класифікатори, як правило, мають кращі значення чутливості, ніж нечіткі системи, і, як наслідок, демонструють кращу здатність виявляти пацієнтів з ІХС, в той час як нечіткі системи досягають відносно кращих результатів у специфічності, що означає кращу достовірність у виявленні пацієнтів без ІХС. Ці два спостереження показують, що роль двох наборів класифікаторів може бути взаємодоповнюючою і може бути використана в стратегії ансамблевого класифікатора для покращення якості НСНП. Результати, досягнуті Д-НСНП (після застосування САК), показують, що покращення чутливості (з 0.750 до 0.792) є більш значним, ніж покращення специфічності (з 0.873 до 0.887). Таким чином, САК загалом покращила здатність НСНП для діагностики ІХС і,

зокрема, для виявлення пацієнта з ІХС, що є бажаною характеристикою для будь-якої системи діагностики ІХС.

Рисунок 4.1 демонструє, як стратегія САК впливає на достовірність системи, надаючи можливість оцінити, наскільки система здатна покращити розпізнавання випадків ІХС.

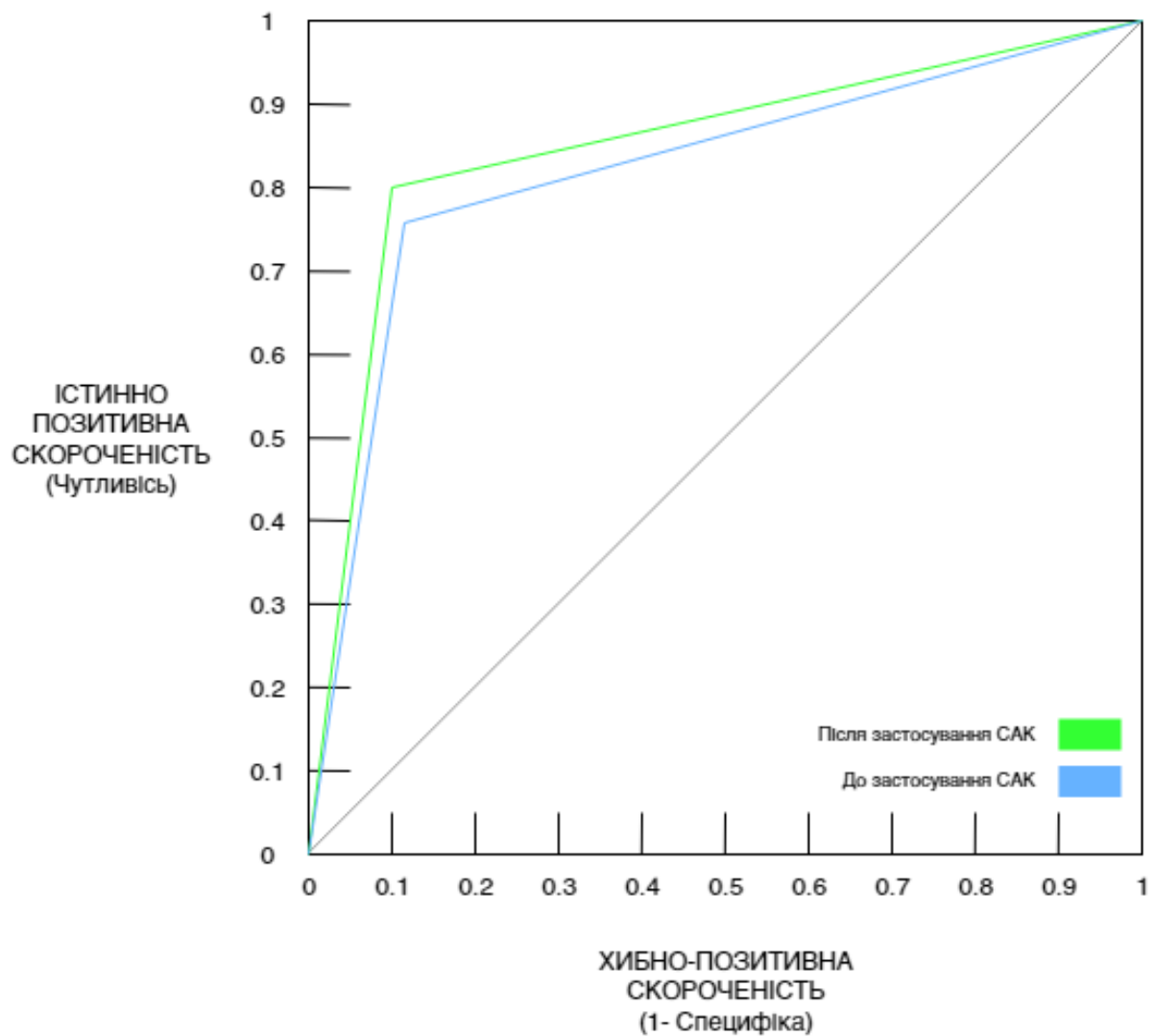


Рисунок 4.1 – ROC-крива НСНП до і після застосування стратегії ансамблевого класифікатора (САК)

Додатково, рисунок 4.2 ілюструє порядок виконання основних етапів процесу діагностики, що реалізовані в НСНП. Ці діаграми є важливим

компонентом для розуміння логіки функціонування системи, а також відображають взаємозв'язок між різними етапами обробки даних та ухвалення рішень.

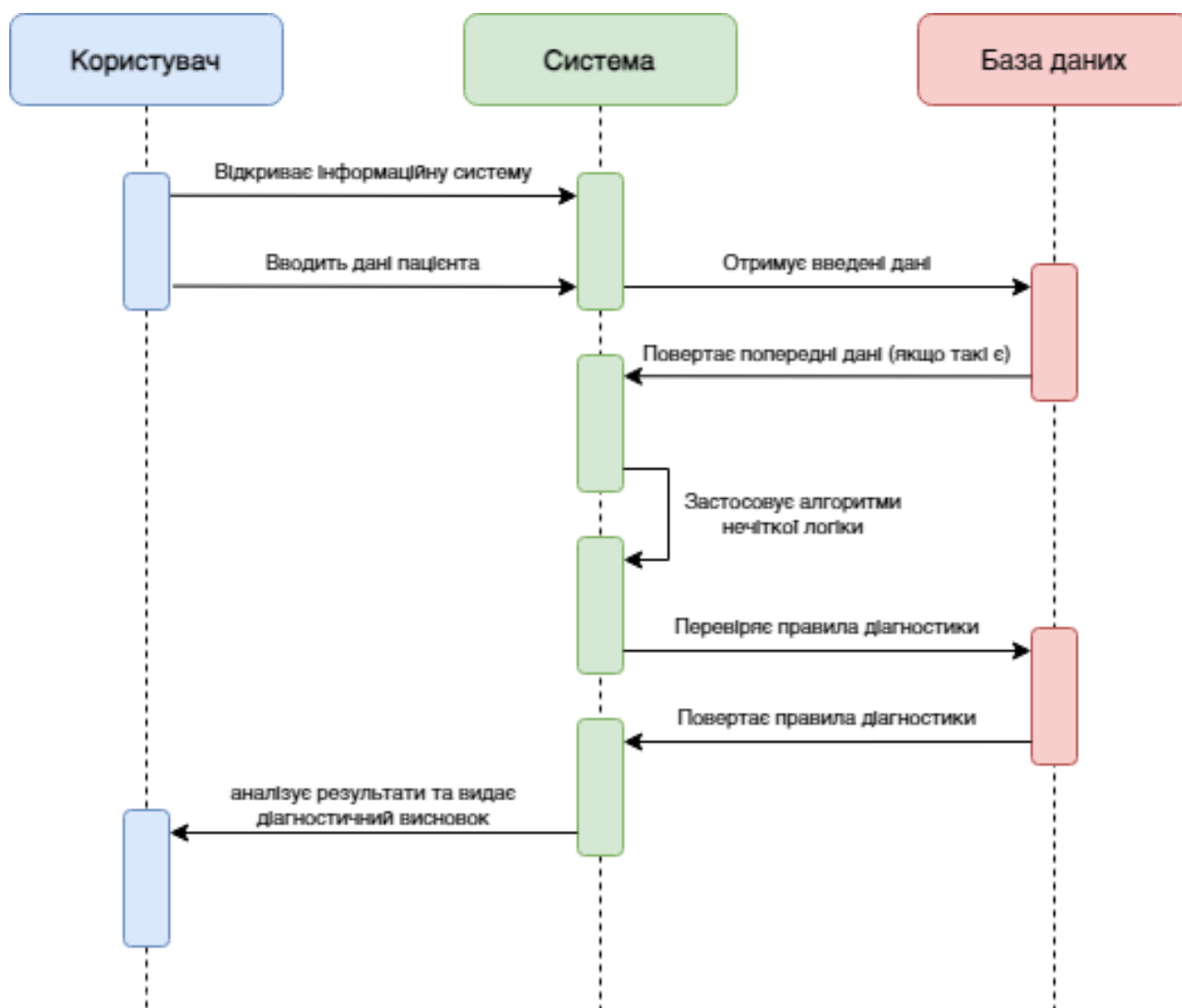


Рисунок 4.2 – Діаграми послідовності

4.4 Лінгвістичні нечіткі правила підмножини 2

Відмова від відповідальності: нечіткі правила, перераховані тут, не повинні використовуватися в клінічній діагностиці без консультації з досвідченими лікарями.

1. ЯКЩО КМС (0), ТОДІ ІХС (відсутня) з вірогідністю: 0.28 та підтримкою: 0.337
2. ЯКЩО ТГБ (безсимптомний) І таласемія (оборотний дефект), ТО ІХС (наявна) з достовірністю: 0,80 та підтримкою: 0.231
3. ЯКЩО стать (чоловіча) І ТГБ (неангінозний біль) І КМС (0) І таласемія (норма), ТОДІ ІХС (відсутня) з достовірністю: 0,72 та підтримкою: 0.124
4. ЯКЩО стать (чоловіча) І ТГБ (безсимптомний), ТОДІ ІХС: (наявна) з достовірністю 0,27 та підтримкою 0.074
5. ЯКЩО стать (Чоловіча) І ТГБ (безсимптомний) І КМС (1) І таласемія (оборотний дефект), ТОДІ ІХС (наявна) з достовірністю: 0,74 і підтримкою: 0.037
6. ЯКЩО ТГБ (неангінозний біль) І КМС (1) І таласемія (оборотний дефект), ТОДІ ІХС (наявна) з вірогідністю: 0,74 та підтримкою: 0.037
7. ЯКЩО стать (жіноча) І ТГБ (типова стенокардія) І КМС (0), І таласемія (норма), ТОДІ ІХС (відсутня) з достовірністю: 1.0 та підтримкою: 0.012
8. ЯКЩО стать (Чоловіча) І КМС (2) І таласемія (норма), ТОДІ ІХС (наявна) з достовірністю: 0,2 та підтримкою: 0.004

4.5 Інтерфейс інформаційної системи для діагностики

На рисунку 4.3 зображено приклад інтерфейсу для збору даних пацієнта та подальшого застосування алгоритмів діагностики. Він складається з полів введення даних, вибіркового елементів та результатів діагностики, відображених на одному вікні.

Цей інтерфейс призначений для зручного та швидкого введення даних для медичної діагностики, з подальшою обробкою даних та виведенням результатів на основі попередньо налаштованого алгоритму. Користувач може

вільно заповнювати інформацію, редагувати її та отримувати результат без необхідності перезавантажувати сторінку або перемикатися на інші вікна.

Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Введення даних пацієнта

Ім'я пацієнта

Вік

Стать Жіноча Чоловіча

Тип болю в грудях

Артеріальний тиск

Рівень холестерину в крові

Результати ЕКГ 0 1 2

Рівень глюкози в крові натще (>120 мг/дл)

Максимальна частота серцевих скорочень

Стенокардія при навантаженні Ні Так

Депресія ST сегмента

Нахил ST сегмента під час фізичних навантажень

Кількість великих судин

Таласемія

Діагностика ішемічної хвороби серця

Результати діагностики

Рисунок 4.3 – Приклад інтерфейсу інформаційної системи

Опис елементів інтерфейсу:

1. Заголовок сторінки: У верхній частині екрану розміщено заголовок "Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу". Він коротко інформує користувача про основну функцію даної системи.

2. Розділ введення даних пацієнта:

- Ім'я пацієнта: Це текстове поле, де користувач може ввести ім'я пацієнта.
- Вік: Текстове поле для введення віку пацієнта.
- Стать: Є два варіанти вибору: Чоловіча або Жіноча. Цей вибір оформлений у вигляді перемикачів.

3. Форма для медичних показників: У цьому розділі користувач вносить ключові параметри здоров'я пацієнта:

- Тип болю в грудях (ТБГ): Вибіркове меню з варіантами
 1. Типова стенокардія
 2. Атипова стенокардія
 3. Неангінозний біль
 4. Безсимптомний
- Артеріальний тиск (АТС): Поле введення значення.
- Рівень холестерину в сироватці крові (РХС): Поле для введення рівня холестерину.
 - Рівень глюкози в крові натще (РГКН): Перемикач для вибору Так або Ні, що відображають факт наявності або відсутності високого рівня глюкози в крові (>120 мг/дл).
 - Результати електрокардіограми (РЕКС): Перемикач з можливістю вибору одного з трьох варіантів:
 - нормальна кардіограма
 - аномальна кардіограма
 - підозріла кардіограма
 - Максимальна частота серцевих скорочень (МЧСС): Поле введення для значень.
 - Стенокардія, яка виникає при фізичних навантаженнях (СПН): Перемикач для вибору варіанту Так або Ні. Цей параметр враховує, чи відчуває пацієнт біль в грудях під час фізичних навантажень.

4. Додаткові медичні показники:

- Депресія ST сегмента, викликана фізичним навантаженням, у порівнянні зі станом спокою (ДСТ): Поле для введення величини, що відображає глибину депресії ST сегмента під час фізичних навантажень.
- Нахил пікового сегмента ST під час фізичного навантаження (НПСТ): Вибіркове меню з трьома варіантами:
 - нахил вгору;

- плоский;
- нахил вниз;
- Кількість магістральних судин (КМС): Поле для введення числа в діапазоні [0, 3], яке представляє кількість великих судин, видимих при флюороскопії.
- Таласемія: Вибіркове меню, що містить такі варіанти:
 - норма;
 - фіксований дефект;
 - оборотний дефект;
- Діагностика ішемічної хвороби серця (ІХС)

4.6 Візуалізація діагностичних даних

У цьому розділі представлені діаграми, які ілюструють основні процеси та функції інформаційної системи підтримки прийняття рішень для визначення медичного діагнозу. Використання діаграм дозволяє наочно продемонструвати структуру системи, взаємодію користувачів з нею, а також процеси оцінки її ефективності. Діаграми, розміщені в цьому розділі, включають:

- Use-Case діаграма – описує функціональні можливості системи з точки зору користувачів.
- Діаграма процесу оцінки ефективності системи – демонструє кроки, що виконуються для оцінки функціональності та стабільності системи.

Ці візуальні матеріали допоможуть краще зрозуміти, як працює система, хто і як взаємодіє з нею, а також яким чином вона оцінюється для покращення своєї ефективності.

На рисунку 4.4 представлено Use-Case діаграму де зображено три основні категорії користувачів системи: Лікар, Пацієнт та Адміністратор. Кожен з цих користувачів виконує різні дії в системі, які відображені у вигляді прецедентів.

Прецеденти показують основні функціональні можливості, доступні для користувачів.

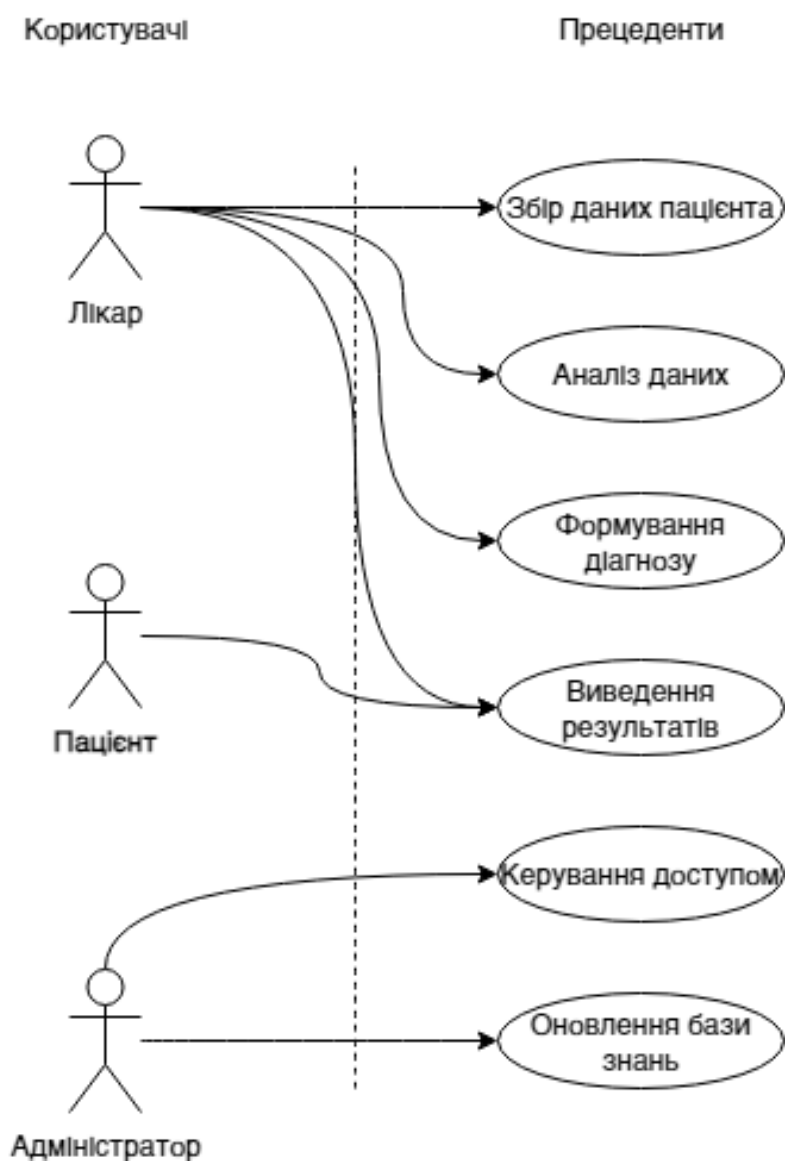


Рисунок 4.4 – Use-Case діаграма

1. Лікар має доступ до таких функцій:

– Збір даних пацієнта – лікар збирає необхідні для діагностики дані від пацієнта.

– Аналіз даних – отримані дані аналізуються з використанням алгоритмів системи для виявлення потенційних захворювань.

- Формування діагнозу – на основі результатів аналізу формується попередній діагноз.

- Виведення результатів – результати діагностики надаються лікарю для подальшого прийняття рішень.

2. Пацієнт взаємодіє із системою через функцію:

- Виведення результатів – пацієнт отримує результати діагностики від системи або від лікаря.

3. Адміністратор відповідає за технічну підтримку та налаштування системи, виконуючи наступні функції:

- Керування доступом – адміністратор регулює права доступу користувачів до системи.

- Оновлення бази знань – оновлює інформацію в базі знань, яка використовується для аналізу даних та формування діагнозу.

Рисунок 4.5 показує процес діаграму оцінки ефективності інформаційної системи на якій зображено послідовність дій для оцінки достовірності, швидкості та стабільності системи.

Процес починається зі збору результатів роботи системи та перевірки її класифікації.

Цей процес можна розглядати як послідовність кроків, що представляють собою логічний і структурований підхід до оцінки ефективності інформаційної системи.

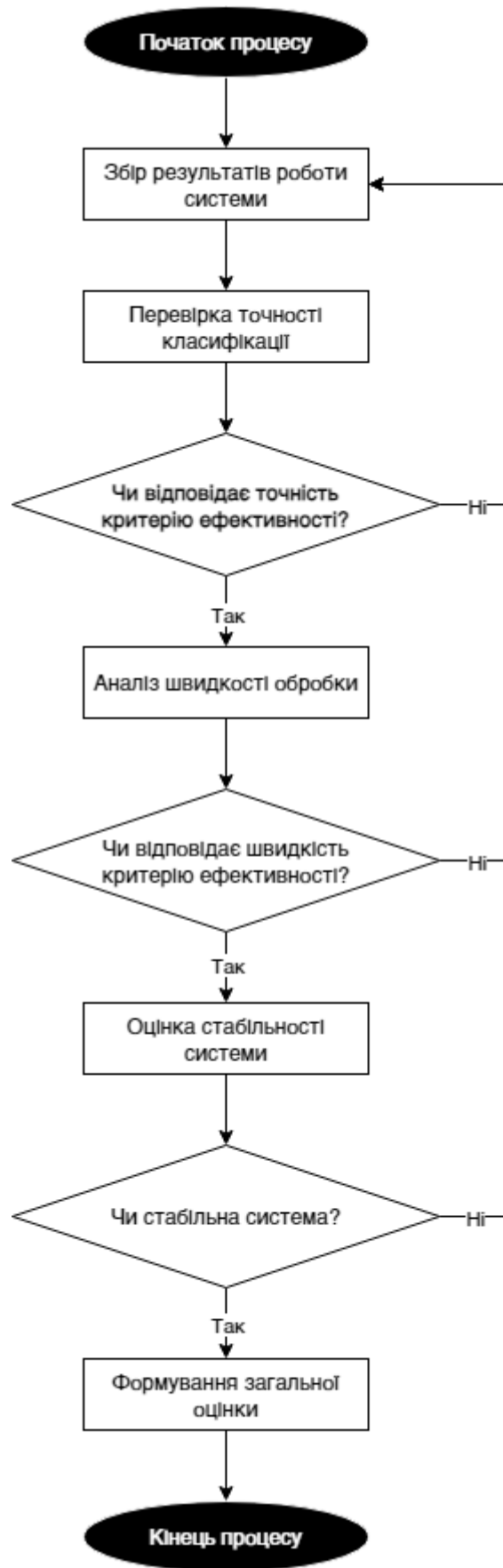


Рисунок 4.5 – Оцінка ефективності Інформаційної системи

Цей процес включає наступні етапи:

1. Збір результатів роботи системи – початковий етап, на якому збираються результати роботи системи для подальшого аналізу.

2. Перевірка достовірності класифікації – аналізується, наскільки достовірно система класифікує дані, відповідаючи заданим критеріям:

– Якщо достовірність не відповідає критеріям, процес повертається на попередній етап для повторного збору даних.

3. Аналіз швидкості обробки – перевіряється, чи відповідає швидкість роботи системи вимогам:

– Якщо швидкість не відповідає критеріям, система повертається до налаштувань.

4. Оцінка стабільності системи – перевіряється стабільність роботи системи:

– Якщо система нестабільна, це повертає процес на етап налаштувань.

5. Формування загальної оцінки – на завершальному етапі формується загальна оцінка ефективності системи, враховуючи достовірність, швидкість та стабільність.

ВИСНОВКИ

У рамках цього дослідження було розроблено систему для діагностики ІХС, засновану на нечітких правилах, яка відзначається високою достовірністю та прозорістю. Нечіткі правила, отримані в результаті роботи системи, формулюються в зрозумілих для лікаря термінах, що описують взаємозв'язки між медичними факторами та ймовірністю діагнозу. Використання природних лінгвістичних термінів дозволяє лікарю легше інтерпретувати інформацію, яку надає система, що значно підвищує її практичну цінність у клінічних умовах.

Однією з головних переваг запропонованої системи є її розширений формат нечітких правил, який, окрім базових характеристик, включає додаткові показники, такі як ступінь достовірності та підтримка кожного правила. Це забезпечує лікаря додатковою інформацією щодо того, наскільки надійне кожне конкретне правило, і чи варто враховувати його під час ухвалення рішення. Таким чином, лікар може не тільки слідувати рекомендаціям системи, але й перевіряти обґрунтованість кожного правила, що покращує контроль над процесом діагностики та знижує ризик помилок.

Крім того, для підвищення достовірності діагностичного висновку в систему інтегровано стратегію ансамблю класифікаторів (САК). Цей підхід використовується у випадках, коли нечітке правило, яке найбільше підходить для діагностики, має низький ступінь визначеності. Стратегія САК дозволяє доповнити діагностичне рішення шляхом залучення декількох додаткових класифікаторів, що підвищує надійність результату. Подібний підхід ґрунтується на концепції медичної практики, де при складних випадках для підтвердження діагнозу можуть залучатися кілька незалежних експертів.

Усі ці особливості роблять розроблену НСНП не лише достовірним інструментом для діагностики ІХС, але й засобом, що дає змогу лікарю глибше розуміти взаємозв'язки між різними факторами ризику та діагнозом ІХС. Це може стати основою для виявлення нових закономірностей і знань у процесі

дослідження ІХС, а також для своєчасного виявлення хвороби на ранніх стадіях, що, у свою чергу, сприятиме зниженню летальності серед пацієнтів.

У першому розділі було проведено аналіз наукової літератури щодо сучасних підходів до діагностики ішемічної хвороби серця (ІХС). Розглянуто особливості використання традиційних методів, які базуються на аналізі клінічних даних, а також їхні обмеження у складних випадках, коли симптоми є нечіткими або маскуються під інші захворювання.

Досліджено перспективи застосування методів машинного навчання, таких як нейронні мережі, метод опорних векторів (МОВ) та нечіткі системи. Було оцінено їхню здатність підвищувати достовірність діагностики, автоматизуючи обробку великих обсягів медичних даних і виявлення прихованих закономірностей. Водночас проаналізовано виклики, зокрема труднощі в інтерпретації результатів і необхідність забезпечення прозорості моделей для їхнього ефективного використання в системах підтримки прийняття рішень.

У другому розділі роботи було сформовано теоретичну основу для процесу діагностики ішемічної хвороби серця (ІХС) із застосуванням сучасних методів машинного навчання. Розглянуто особливості моделей, таких як метод опорних векторів (МОВ), нейронні мережі та нечіткі системи, із визначенням їхніх переваг та обмежень у контексті медичної діагностики. Також запропоновано модель процесу діагностики ІХС, яка охоплює етапи збору, обробки даних, побудови та оцінки класифікаторів.

Особливу увагу приділено стратегії використання ансамблю класифікаторів, яка дозволяє об'єднати результати кількох моделей для підвищення достовірності діагностики.

У третьому розділі було розроблено структуру та визначено ключові компоненти. Основна увага приділялась модулю збору та стандартизації медичних даних, отриманих із різних діагностичних пристроїв, а також модулю обробки та класифікації цих даних.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Boukezzi L., Mesbah S., Bouamama S. A new approach for medical diagnosis based on fuzzy logic and expert systems. *International Journal of Computer Applications*. 2021. Vol. 975. P. 45–55.
2. Uddin M. Z. A wearable sensor-based activity prediction system to facilitate edge computing in smart healthcare system. *Journal of Parallel and Distributed Computing*. 2019. Vol. 123. P. 46–53.
3. Zahan S. A fuzzy approach to computer-assisted myocardial ischemia diagnosis. *Artificial Intelligence in Medicine*. 2001. Vol. 2. P. 271–275.
4. Smith J., Wang Y., Patel S. Hybrid model for cardiovascular risk prediction combining logistic regression and SVM. *Computers in Biology and Medicine*. 2018. P. 28–34.
5. Patel R., Kumar M., Singh A. Deep learning for coronary artery angiography analysis. *IEEE Transactions on Medical Imaging*. 2017. P. 100–107.
6. Garcia L., Torres D., Martinez F. Biomarker-based system for early diagnosis of IHD using KNN and decision trees. *Health Informatics Journal*. 2021. P. 67–75.
7. Nakamura T., Suzuki H., Ito K. LSTM-based model for IHD risk prediction. *Journal of Clinical Cardiology*. 2022. P. 90–95.
8. Roberts M., et al. Using logistic regression and neural networks for risk factor analysis in IHD. *Journal of Clinical Analytics*. 2020. P. 145–153.
9. Oliveira P., Mendes J., Goncalves B. Mobile decision-support platform for clinical use in IHD diagnosis. *Telemedicine Journal*. 2019. P. 56–62.
10. Chang L., Kim D., Liu Y. Adaptive diagnostic system for personalized IHD assessment. *Medical Informatics*. 2021. P. 78–85.
11. Lee Y., Park H., Kim J. ECG-based risk prediction platform for IHD with deep learning integration. *Computers in Cardiology*. 2018. P. 213–218.
12. Gupta S., Roy R., Debnath P. Fuzzy logic system for early diagnosis of Parkinson's disease. *Journal of Mult. Valued Logic & Soft Computing*. 2020. Vol. 34. P. 109–126.

13. Mohammed S., Attal F., Amirat Y. A wearable sensor-based fuzzy system for gait analysis and rehabilitation monitoring. *Sensors*. 2018. P. 21.
14. Zadeh L.A. Fuzzy sets. *Information and Control*. 1965. Vol. 8. P. 338–353.
15. Jain A., Dubey S. An expert system approach for diagnosis of chronic diseases using fuzzy logic. *IEEE Transactions on Biomedical Engineering*. 2019. P. 251–260.
16. Rassel R., Zhang X., Li Y. A fuzzy logic-based clinical decision support system for diabetes management. *Journal of Medical Informatics*. 2017. P. 14–25.
17. Ordóñez F., Toledo P., Sanchis A. Hybrid models for human activity recognition using binary sensors. *Sensors*. 2017. P. 5460–5477.
18. Chen X., Wang L. Application of fuzzy logic to intelligent control in medical devices. *Control and Automation Systems*. 2018. Vol. 22. P. 322–337.
19. Zou Y., Xiao J., Han J. Gesture recognition system using RFID-based sensors. *IEEE Transactions on Mobile Computing*. 2020. P. 381–393.
20. Pradhan S., Chai E. Pervasive touch interface using RFID. *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. Snowbird, UT, USA, 16–20 October 2017. P. 261–274.
21. Tan T., Lee J. Fuzzy-based predictive models for disease outbreaks. *AAAI Conference on Artificial Intelligence*. Phoenix, USA, 12–17 February 2019. P. 194–200.
22. Schmidhuber J. Long short-term memory (LSTM) networks. *Neural Computation*. 2019. Vol. 32.
23. Rajaraman A., Ullman J.D. *Mining of Massive Datasets*. Cambridge : Cambridge University Press, 2019.
24. Du Y., Li Z. Progressive transmission for visual descriptors using saliency. *Journal of Mult. Valued Logic and Soft Computing*. 2018. P. 125–145.
25. Gomaa W.H., Fahmy A.A. A fuzzy logic-based system for human emotions recognition. *Procedia Computer Science*. 2020. Vol. 177. P. 123–134.
26. Hochreiter S., Schmidhuber J. Advanced applications of LSTM in health informatics. *IEEE Transactions on Neural Networks*. 2019. Vol. 30.

27. Du Y., Li Z. Mobile visual search techniques based on saliency. *Mobile Vision*. 2013. P. 922–928.
28. Attal F., Chamroukhi F., Oukhellou L. Physical activity monitoring using hybrid sensors. *Sensors*. 2018. P. 15.
29. Gupta S., Debnath P. Development of expert systems for chronic disease monitoring. *IEEE Biomedical Engineering*. 2019. Vol. 29. P. 221–233.
30. Voulodimos A., Doulamis N., Doulamis A., Protopapadakis E. Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*. 2018. Article ID 7068349. P. 13.
31. Park H., Lee S., Kim K. Convolutional neural network for biomedical image segmentation. *IEEE Transactions on Medical Imaging*. 2019. 38(7). P. 1539–1548.
32. Lee J., Kim D., Park Y. An ensemble-based approach for predicting cardiovascular events. *Journal of Medical Systems*. 2021. 45(3). P. 33.
33. Sun Y., Li X., Liu Z. Fuzzy logic applications in healthcare: A systematic review. *IEEE Reviews in Biomedical Engineering*. 2020. 13. P. 76–95.
34. Khatibi V., Montazer G.A. A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment. *Expert Systems with Applications*. 2010. 37(12). P. 8536–8542.
35. Tan J., Zhang Y., Zhang X. Predictive models in cardiology: Comparing deep learning and traditional approaches. *Journal of Clinical Cardiology*. 2019. 85(12). P. 145–152.
36. Afonso V., Tompkins W.J., Nguyen T.Q., Luo S. ECG beat detection using filter banks. *IEEE Transactions on Biomedical Engineering*. 1999. 46(2). P. 192–202.
37. Singh D., Choudhury S. Machine learning algorithms for early detection of IHD. *International Journal of Healthcare Informatics*. 2020. 11(1). P. 45–52.
38. Martis R.J., Acharya U.R., Min L.C., Suri J.S. ECG analysis based on wavelet transform and fuzzy logic. *Biomedical Signal Processing and Control*. 2013. 8(6). P. 673–680.

39. Abdar M., Yazdanpanah S., Tahmasbi S. Performance analysis of machine learning techniques in diagnosis of chronic heart diseases. *Journal of Medical Systems*. 2019. 43(8). P. 58.
40. Ibaida A., Khalil I. Multi-sensor data fusion in wearable sensor networks for cardiac health monitoring. *Information Fusion*. 2016. 33. P. 91–102.
41. Sharma A., Verbeke W. Improving diagnosis of IHD with deep learning. *Computer Methods and Programs in Biomedicine*. 2019. 177. P. 91–98.
42. Alim U., Latif G., Ali S. Hybrid deep learning models for detection of cardiovascular diseases. *Applied Soft Computing*. 2021. 110. Article ID 107588.
43. Mohanty S., Panda S. A fuzzy expert system for diagnosis of cardiovascular diseases. *Procedia Computer Science*. 2018. 132. P. 290–297.
44. Ghorbani A., Heidarysafa M., Qaderi M., Rahmani H. Comprehensive study on deep learning models for cardiovascular disease diagnosis. *Journal of Biomedical Informatics*. 2021. 119. Article ID 103815.
45. Wang Z., Wang W. Ensemble learning in medical diagnosis of heart diseases. *Medical & Biological Engineering & Computing*. 2018. 56(11). P. 1987–1998.
46. Kamran S., Alharbi A. ECG classification using hybrid LSTM model. *Computers in Biology and Medicine*. 2021. 133. Article ID 104413.
47. Haghani S., Altunkaya K., Hadavandi E. Coronary artery disease prediction based on machine learning methods. *Expert Systems with Applications*. 2021. 164. Article ID 113801.
48. Tadesse G., Mehtsun S., Xiong L. Comparative analysis of different machine learning techniques for cardiac disease classification. *International Journal of Medical Informatics*. 2020. 139. Article ID 104124.
49. Chaurasia V., Pal S. Early detection of heart diseases using data mining. *Asian Journal of Computer Science and Information Technology*. 2013. 2(2). P. 68–71.
50. Cai H., Xu T., Zhang Y. Machine learning in heart disease prediction: Recent trends and challenges. *Journal of Medical Systems*. 2020. 44(5). P. 92.
51. Zhang X., Yu X., Zhao Y. Smart healthcare: Big data and artificial intelligence applications. *Future Generation Computer Systems*. 2019. 108. P. 683–698.

52. El-Sappagh S., Ali F., Islam S.M.R. A review of machine learning in healthcare: Opportunities, challenges, and applications. *IEEE Access*. 2021. 9. P. 39093–39134.
53. Kumar S., Gupta A., Sharma N. Decision support systems for cardiovascular diagnosis using hybrid methods. *Journal of King Saud University-Computer and Information Sciences*. 2021. 33(6). P. 602–614.
54. Wu X., Chen S., Zhang J. Application of machine learning in cardiovascular disease prediction. *Frontiers in Cardiovascular Medicine*. 2020. 7. Article ID 234.
55. Roy S., Mukherjee M. Hybrid models in cardiology: Combining data-driven and mechanistic models. *Medical Physics*. 2021. 48(2). P. 191–200.
56. Wang H., Yan L., Xu Y. A deep learning approach for ECG-based coronary artery disease diagnosis. *Computers in Biology and Medicine*. 2021. 131. Article ID 104235.
57. Fekri-Ershad S. Applying deep neural networks in wearable devices for heart monitoring. *IEEE Transactions on Biomedical Circuits and Systems*. 2020. 14(2). P. 308–319.
58. Saxena S., Verma K., Thakur D. Comparative study of various machine learning techniques for cardiovascular risk prediction. *Journal of Computer Science and Technology*. 2021. 25(3). P. 94–101.
59. Li T., Liao H., Wan Z. Development of fuzzy logic-based diagnostic tools for heart disease. *Journal of Intelligent and Fuzzy Systems*. 2020. 39(6). P. 7851–7862.
60. Chen J., Zhao L., Li K. Applications of machine learning in cardiovascular risk assessment. *IEEE Access*. 2020. 8. P. 87541–87556.
61. Yue P., Wang F., Liu J. Fuzzy logic for decision support in clinical medicine. *Artificial Intelligence in Medicine*. 2018. 90. P. 34–42.
62. Qian H., Lin T., Chen H. Predicting IHD risk with machine learning and data mining techniques. *Applied Soft Computing*. 2021. 105. Article ID 107258.

63. Mittal N., Rana K., Singh D. Internet of Medical Things (IoMT): A new frontier for healthcare. *Computers in Biology and Medicine*. 2019. 106. P. 160–170.
64. Mishra R., Jha S., Verma P. Fuzzy rule-based expert systems for heart disease diagnosis. *Procedia Computer Science*. 2020. 171. P. 1067–1075.
65. Pławiak P., Abdar M., Acharya U.R. Data fusion and ensemble learning approaches for multi-sensor-based cardiovascular diagnosis. *Sensors*. 2019. 19(18). Article ID 3964.
66. Xia Y., Wu Y., Yin S. Machine learning applications for cardiovascular health. *IEEE Reviews in Biomedical Engineering*. 2021. 14. P. 77–92.
67. Sinha A., Kumar D. Real-time monitoring and diagnosis of heart diseases using machine learning. *Journal of Clinical Monitoring and Computing*. 2020. 34(4). P. 703–710.
68. Arora R., Pandey S. Deep learning applications in clinical cardiology. *Computers in Biology and Medicine*. 2021. 137. Article ID 104778.
69. Khan Z., Butt F.M., Malik J. Comparative analysis of machine learning classifiers for cardiovascular diagnosis. *Computers in Biology and Medicine*. 2019. 116. Article ID 103387.
70. Chen T., He T., Zhou B. LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*. 2017. Vol. 30. P. 3146–3154.
71. Bharti P., Gole P., Mehra A. An efficient heart disease prediction model based on machine learning techniques. *Informatics in Medicine Unlocked*. 2021. 25. Article ID 100681.
72. Kaur M., Kaur S., Kaur K. Automated diagnostic system for cardiac diseases using ensemble learning techniques. *Expert Systems with Applications*. 2020. 143. Article ID 113071.
73. Sharma A., Jain V. Intelligent decision support systems for chronic disease management. *IEEE Journal of Biomedical and Health Informatics*. 2021. 25(2). P. 371–378.

74. Wang C., Sun X., Zhang L. Hybrid models for cardiovascular risk prediction in elderly patients. *International Journal of Medical Informatics*. 2019. 132. Article ID 103987.

ДОДАТОК А

КОПІЯ ОПУБЛІКОВАНОЇ НАУКОВОЇ СТАТТІ

УДК 004.052.2

Кульбачний В.В., Данчук С.В., Нічепорук А.О.

Хмельницький національний університет

ІНФОРМАЦІЙНА СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ

У роботі розглядаються прикладні аспекти розробки інформаційної системи для автоматизації процесів збору та аналізу даних пацієнтів у медичних закладах. Розроблена система забезпечує точне формування діагнозів, оперативне виведення результатів та ефективне керування доступом до даних. Вона також включає механізм оновлення бази знань, що дозволяє підтримувати її актуальність і адаптованість до нових медичних досліджень. Рішення спрямоване на покращення швидкості та точності прийняття рішень у медичній практиці та підвищення якості обслуговування пацієнтів.

This work examines the applied aspects of developing an information system for automating data collection and analysis processes in healthcare institutions. The proposed system ensures precise diagnosis generation, rapid results display, and effective data access management. It also includes a knowledge base update mechanism to maintain relevance and adaptability to new medical research. The solution aims to improve the speed and accuracy of decision-making in medical practice and enhance the quality of patient care.

У сучасному суспільстві інформаційні технології відіграють ключову роль у всіх сферах, зокрема в охороні здоров'я. Проблема управління даними пацієнтів стає особливо актуальною, оскільки збільшення обсягу інформації вимагає нових підходів до її обробки та аналізу. Недостатня автоматизація медичних процесів призводить до затримок у наданні допомоги, а також підвищує ризик помилок при діагностиці. В умовах, коли кожна хвилина може бути критично важливою, ефективність інформаційних систем стає вирішальним фактором для забезпечення високої якості медичних послуг.

Крім того, важливою проблемою залишається недостовірність і неповнота медичних даних, які можуть суттєво вплинути на прийняття рішень лікарями. Це не тільки погіршує якість діагностики, але й підвищує ризик помилок у лікуванні.

Сучасні технології, зокрема машинне навчання, можуть стати важливими інструментами для покращення точності аналізу медичних даних і формування більш обґрунтованих діагнозів.

Аналіз існуючих рішень у цій сфері вказує на їх численні недоліки, такі як складність інтеграції з іншими системами та висока вартість впровадження. Багато медичних установ не мають можливості дозволити собі сучасні технології через фінансові обмеження, що, у свою чергу, заважає їм забезпечити своєчасне і якісне обслуговування пацієнтів. Ця ситуація потребує термінового вирішення шляхом розробки нових технологій, які були б доступнішими і більш ефективними в умовах сучасної медицини.

Невирішені питання щодо навчання медичного персоналу для роботи з новими інформаційними системами також заслуговують на увагу. Сучасні технології потребують не лише технічних знань, а й здатності адаптуватися до нових умов праці. Помилки, які виникають через недостатню кваліфікацію, можуть мати серйозні наслідки для пацієнтів. Тому важливо враховувати не лише технологічні аспекти, а й людський фактор, що впливає на успішність впровадження нових рішень у медичній практиці.

Зважаючи на ці виклики, дослідження, присвячене розробці інформаційних систем для автоматизації медичних процесів, є надзвичайно актуальним. Воно дозволить виявити нові методи і рішення, які зможуть поліпшити обслуговування пацієнтів, зменшити ймовірність помилок і підвищити загальну ефективність медичних установ. Використання таких технологій може стати ключем до забезпечення не лише високої якості медичних послуг, а й покращення загального стану здоров'я населення.

У цьому контексті, важливо визначити не лише технічні вимоги до інформаційних систем, а й розробити стратегії їхнього впровадження, навчання медичного персоналу та підтримки користувачів. Це дозволить з максимальною ефективністю використовувати можливості сучасних технологій у медицині, створюючи умови для їхнього швидкого та успішного впровадження в реальну практику.

Упродовж останніх років було проведено численні дослідження, присвячені автоматизації медичних процесів та застосуванню інформаційних технологій у медицині. Серед актуальних напрямків можна виділити використання електронних медичних записів (ЕМЗ), систем підтримки прийняття рішень (СПП), а також технології телемедицини. ЕМЗ значно полегшують

зберігання та обробку медичної інформації, знижуючи ризики помилок, пов'язаних із паперовою документацією.

Публікації, що аналізують ефективність СПП, демонструють, що такі системи можуть значно покращити точність діагностики. Зокрема, роботи, присвячені алгоритмам машинного навчання, показують, що їх застосування в медичній практиці дозволяє швидше та ефективніше виявляти патології, що позитивно впливає на результати лікування пацієнтів.

Останні дослідження в галузі медицини та інформаційних технологій акцентують увагу на застосуванні алгоритмів машинного навчання для автоматизації обробки медичних даних. Зокрема, в публікаціях відзначаються досягнення в розробці методів, що дозволяють знижувати ймовірність помилок у діагностиці завдяки аналізу великих обсягів даних. Наприклад, вивчені аспекти глибокого навчання виявляються ефективними для обробки зображень, що відкриває нові можливості для покращення діагностики різних захворювань.

Незважаючи на ці досягнення, залишається ряд невирішених проблем, зокрема потреба в адаптації існуючих рішень до специфічних умов роботи різних медичних установ. Багато з поточних моделей виявляються недостатньо гнучкими і не можуть ефективно працювати з даними, отриманими в умовах реальної клінічної практики. Це свідчить про необхідність подальших досліджень у цьому напрямку для оптимізації алгоритмів і підвищення їх ефективності.

Однак, попри прогрес у цій галузі, існує чимало невирішених питань. Наприклад, недостатня увага до інтеграції нових технологій з уже існуючими системами може призвести до низької ефективності їх використання. Також важливим є питання безпеки та конфіденційності даних пацієнтів, оскільки зростання обсягу інформації, що обробляється, підвищує ризики витоку чутливих даних.

Дослідження також вказують на необхідність розвитку стандартів для взаємодії між різними інформаційними системами. Це дозволить створити ефективну інфраструктуру для обміну даними між установами охорони здоров'я, що сприятиме підвищенню якості медичних послуг. Невирішеними залишаються питання адаптації нових технологій до потреб медичного персоналу та пацієнтів, що вимагає детального аналізу та подальших досліджень у цій сфері.

Таким чином, основні публікації, на які спирається автор, підкреслюють як досягнення, так і виклики у впровадженні інформаційних технологій у медичну практику. Визначені невирішені проблеми свідчать про необхідність подальших досліджень, які

сприятимуть створенню ефективних та безпечних інформаційних систем для покращення якості медичних послуг.

Метою даного дослідження є розробка ефективних інформаційних технологій для автоматизації процесів в медичній сфері, зокрема для поліпшення діагностики та лікування пацієнтів. Це передбачає створення системи, яка інтегрує існуючі медичні дані, алгоритми аналізу та методи візуалізації, що забезпечить своєчасну та точну інформацію для медичного персоналу.

Для досягнення цієї мети необхідно розглянути впровадження нових технологій, які дозволять зібрати, обробити та проаналізувати медичні дані пацієнтів. Це включає дослідження можливостей використання штучного інтелекту та алгоритмів машинного навчання для підвищення точності діагностики. Основною задачею є оцінка ефективності впроваджених рішень на базі реальних даних, що дозволить виявити сильні та слабкі сторони запропонованої системи.

Важливо також врахувати специфіку медичного контексту, в якому працюватиме інформаційна система. Це передбачає аналіз потреб медичних працівників та пацієнтів, щоб забезпечити зручність користування системою. Мета дослідження включає не лише технічні аспекти, але й соціальні фактори, які можуть вплинути на прийняття нових технологій у медицині.

Завданням дослідження є також вивчення впливу соціальних і екологічних факторів на здоров'я пацієнтів. Це дозволить створити більш комплексні моделі, які враховують різні аспекти життя людини, що впливають на стан її здоров'я. Виконання цих завдань сприятиме підвищенню точності діагностики та якості медичних послуг.

Таким чином, дослідження спрямоване на створення моделі, яка б відповідала сучасним вимогам медичної практики, дозволяючи підвищити ефективність лікування та покращити якість надання медичних послуг. Результати роботи будуть сприяти розвитку інноваційних рішень у галузі охорони здоров'я та інтеграції нових технологій у повсякденну практику.

У рамках дослідження були розроблені основні матеріали, що включають в себе алгоритми для обробки та аналізу медичних даних. Основна увага приділяється використанню методів машинного навчання для створення моделей, які здатні автоматично класифікувати медичну інформацію, виявляти аномалії та пропонувати діагнози на основі історії захворювань пацієнтів. Відзначено, що такі технології вже довели свою ефективність у ряді досліджень, і їх адаптація для специфіки нашої медичної системи може суттєво підвищити рівень обслуговування пацієнтів.

Дослідження також включає порівняння існуючих інформаційних систем для управління медичними даними, що дозволяє виявити недоліки та прогалини в їх функціонуванні. На основі цього аналізу були визначені критерії, яким повинна відповідати нова система: простота використання, швидкість обробки даних, можливість інтеграції з іншими медичними платформами, а також безпека даних пацієнтів. Розробка концепції нової системи базується на аналізі потреб користувачів, зокрема медичних працівників і пацієнтів.

Крім того, було проведено тестування запропонованих алгоритмів на реальних наборах даних, що дало змогу оцінити їх ефективність та точність. Результати тестування підтвердили, що система здатна не лише покращити точність діагностики, а й зменшити час, необхідний для обробки інформації. На підставі отриманих результатів розроблено рекомендації щодо впровадження нових методів у медичну практику, що забезпечить їх ефективне використання.

Особливу увагу було приділено аспектам безпеки даних, адже зберігання та обробка медичної інформації вимагають дотримання суворих стандартів конфіденційності. Запропоновані заходи безпеки включають шифрування даних, доступ на основі ролей та регулярний моніторинг системи на предмет вразливостей. Ці елементи забезпечать захист особистої інформації пацієнтів та збереження довіри до нових технологій.

Таким чином, основні матеріали дослідження підкреслюють важливість інтеграції сучасних інформаційних технологій у медичну сферу, сприяючи покращенню діагностики та лікування пацієнтів. Результати дослідження демонструють, що інноваційні підходи здатні суттєво підвищити якість медичних послуг та оптимізувати роботу медичних закладів.

У результаті проведеного дослідження встановлено, що впровадження нових інформаційних технологій у медичну практику має потенціал для значного покращення якості обслуговування пацієнтів та підвищення точності діагностики. Основні результати підтверджують ефективність розроблених алгоритмів машинного навчання для автоматизації процесів обробки медичних даних, а також виявлення аномалій у стані здоров'я пацієнтів. Це забезпечить своєчасне прийняття рішень лікарями на основі об'єктивної інформації, зменшуючи ризики помилок у діагностиці.

Крім того, у ході аналізу існуючих рішень виявлено, що сучасні інформаційні системи не завжди відповідають потребам користувачів, зокрема лікарів і пацієнтів. Це свідчить про необхідність подальшої

розробки адаптованих рішень, які зможуть задовольнити специфічні вимоги різних медичних установ. Важливо також акцентувати увагу на інтеграції нових технологій із вже існуючими системами для забезпечення безперебійного доступу до медичних даних та зручності їх використання.

Перспективи подальших досліджень у даному напрямку включають розробку більш складних моделей, які враховують різні параметри, такі як генетична інформація та поведінкові фактори пацієнтів. Вивчення впливу соціальних і екологічних чинників на здоров'я також може відкрити нові горизонти для вдосконалення медичних технологій. Додатково, подальші дослідження можуть зосередитися на оцінці ефективності запроваджених технологій у клінічній практиці та вивченні їх впливу на результати лікування пацієнтів.

Отже, підсумовуючи, можна стверджувати, що інноваційні технології в медицині є перспективним напрямком, що відкриває нові можливості для покращення якості медичних послуг. Наступні етапи дослідження стануть ключовими для впровадження нових рішень у практику, що дозволить забезпечити більш ефективну та безпечну систему охорони здоров'я.

Перелік посилань

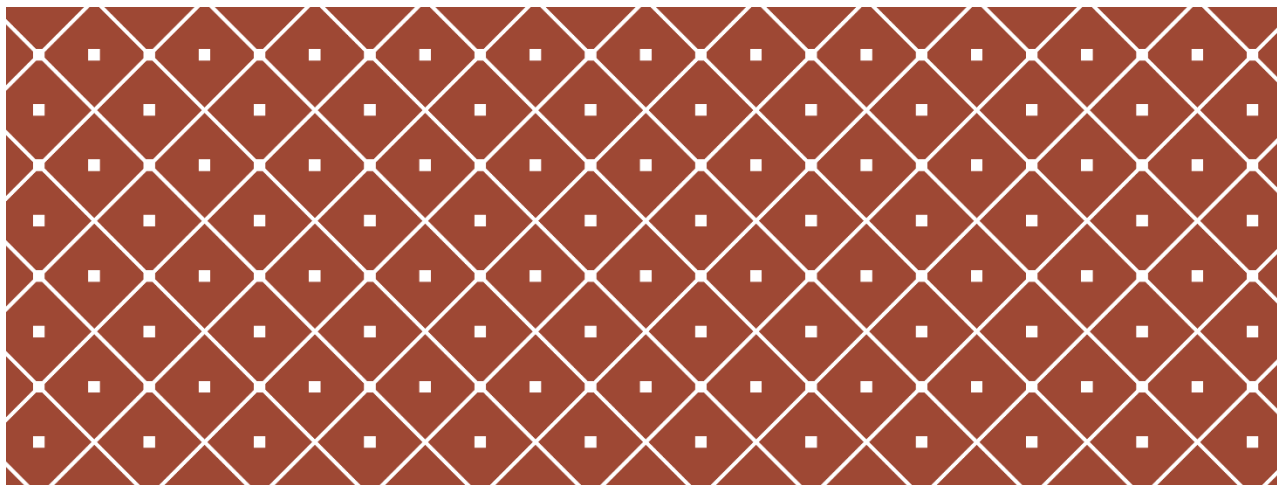
1. Ordóñez, F.; Toledo, P.; Sanchis, A. Hybrid models for human activity recognition using binary sensors. *Sensors* 2017, 5460–5477 pp.
2. Chen, X.; Wang, L. Application of fuzzy logic to intelligent control in medical devices. *Control and Automation Systems* 2018, Vol. 22, 322–337 pp.
3. Zou, Y.; Xiao, J.; Han, J. Gesture recognition system using RFID-based sensors. *IEEE Transactions on Mobile Computing* 2020, 381–393 pp.

Дані про авторів (не для друку):

ПІБ автора	Телефон	Email
Кульбачний Владислав Васильович	+380965654069	shup208@gmail.com
Нічепорук Андрій Олександрович	+380964687613	andrey.nicheporuk@gmail.com
Данчук Сергій Вікторович	+380977264253	sergey.danchuk.p@gmail.com

ДОДАТОК Б

ПРЕЗЕНТАЦІЯ ДО ЗАХИСТУ



ІНФОРМАЦІЙНА СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ МЕДИЧНОГО ДІАГНОЗУ

Кульбачний В.В.

Науковий керівник к.т.н, доцент
Нічепорук А.О.

Хмельницький, 2024

Об'єктом дослідження є процес діагностики медичних станів з використанням інформаційних систем підтримки прийняття рішень.

Предметом дослідження є методи і моделі машинного навчання, які використовуються для розробки системи підтримки прийняття рішення, спрямованої на точне встановлення медичного діагнозу.

Метою кваліфікаційної роботи магістра є підвищення достовірності формування медичного діагнозу шляхом розробки інформаційної системи підтримки прийняття рішення для діагностики медичних станів

НАУКОВА НОВИЗНА

1. Набула подальшого розвитку інформаційна система підтримки прийняття рішення для визначення медичного діагнозу, яка на відміну від відомих залучає адаптивні алгоритми машинного навчання для аналізу великих обсягів медичних даних, що дозволило підвищити достовірність діагностичних процесів
2. Набула подальшого розвитку модель процесу генерації нечітких правил, яка, на відміну від відомих, включає значення підтримки правил та їх ступінь достовірності, що дозволило враховувати не лише точність класифікації, а й частоту виникнення правил у навчальних даних, що зменшило ймовірність помилкових узагальнень.

ПОСТАНОВКА ЗАДАЧІ

Постановка задачі дослідження полягає в розробці інформаційної системи підтримки прийняття рішень для діагностики ішемічної хвороби серця, яка використовує методи машинного навчання та штучного інтелекту для аналізу медичних даних, таких як електрокардіограми та лабораторні аналізи. Метою є автоматизація процесу обробки даних і підтримка лікарів у прийнятті достовірних і швидких діагностичних рішень, що має значно знизити ризики хибнопозитивних та хибнонегативних результатів.

Завдання дослідження включають аналіз існуючих методів діагностики ІХС, розробку архітектури інформаційної системи, її реалізацію та тестування на реальних медичних даних, а також порівняльний аналіз результатів роботи системи з традиційними методами діагностики. Окрім технічних аспектів, важливо оцінити вплив цієї системи на клінічний процес, зокрема на час діагностики, достовірність діагнозу та зниження кількості помилкових рішень.

Очікується, що створений прототип системи не тільки продемонструє високу достовірність у класифікації стану пацієнтів, але й забезпечить зручний інтерфейс для лікарів, що сприятиме швидкій адаптації в клінічній практиці.

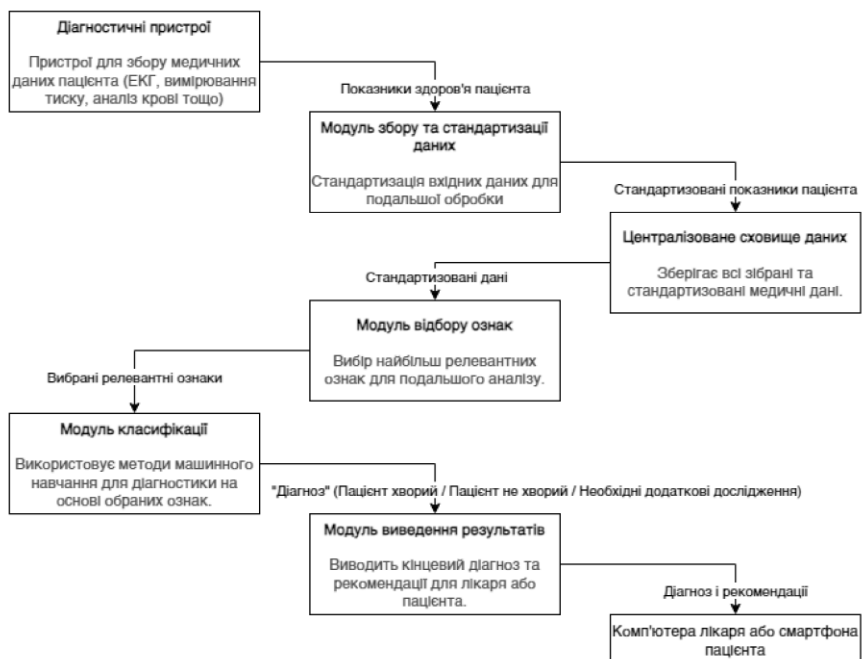
ЕТАПИ РОЗРОБКИ

Основні етапи розробки системи підтримки прийняття рішення для діагностики медичних станів включають:

1. Збір та підготовка даних
2. Відбір ознак та попередня обробка
3. Розробка моделі машинного навчання
4. Оцінка якості та тестування
5. Інтеграція в інформаційну систему



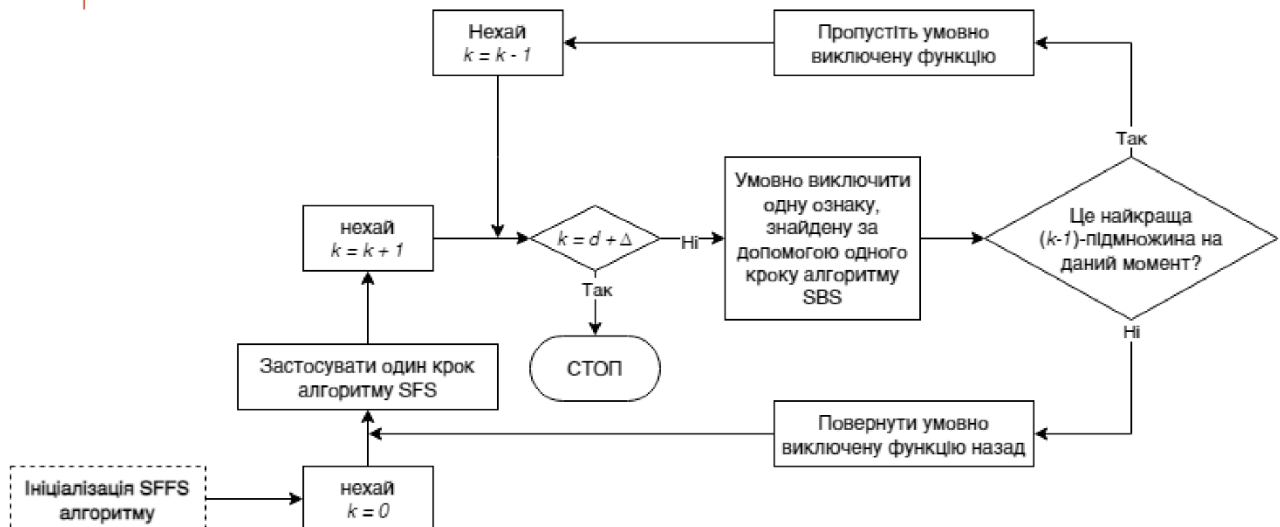
СТРУКТУРА ІС



ЗБІР ТА ПІДГОТОВКА ДАНИХ

Атрибути	Опис
Вік	Вік
Стать	Стать
ТГБ	Тип болю в грудях
АТС	Артеріальний тиск у стані спокою
РХС	Рівень холестерину в сироватці крові (мг/дл)
РГКН	Рівень глюкози в крові натще (>120 мг/дл)
РЕКС	Результати електрокардіограми в стані спокою
МЧСС	Максимальна досягнута частота серцевих скорочень
СФН	Стенокардія, викликана фізичним навантаженням
ДСТ	Депресія ST сегмента.
НПСТ	Нахил пікового сегмента ST під час фізичного навантаження
КМС	Кількість магістральних судин
Таласемія	Таласемія
ІХС	Діагностика ішемічної хвороби серця

АЛГОРИТМ SFFS



РОЗРОБКА МОДЕЛІ МАШИННОГО НАВЧАННЯ

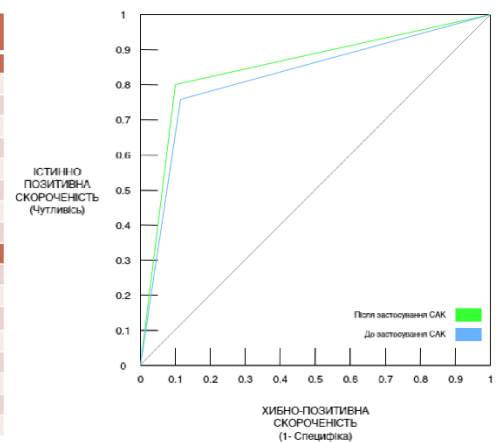
Процедура починається з відбору релевантних ознак методом SFFS. Далі формується набір нечітких правил, кожне з яких має формат "якщо-то", що дозволяє врахувати комбінації обраних ознак для класифікації діагнозу. Оцінка сумісності навчального прикладу $a_t = (a_{t1}, a_{t2}, \dots, a_{tn})$ з правилом R_k обчислюється як:

$$\mu_{R_k}(a_t) = \mu_{k1}(a_{t1}) * \mu_{k2}(a_{t2}) * \dots * \mu_{kn}(a_{tn})$$

де $\mu_{k_i}(\cdot)$ — функція належності для нечіткої множини χ_j^k . Після цього кожне правило оптимізується за допомогою алгоритму NSGA-II для підвищення достовірності та зменшення кількості правил.

ОЦІНКА ЯКОСТІ ТА ТЕСТУВАННЯ

Методи	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	В середньому
Навчальні дані											
С4.5	86.80	85.10	86.40	86.40	86.00	86.80	86.80	85.10	87.20	86.40	86.30
ШНМ	87.60	85.50	86.80	87.20	86.00	87.20	87.60	85.50	87.60	87.20	86.80
HF-GBML	86.83	85.19	85.19	86.42	86.01	86.01	87.24	85.60	86.83	86.83	86.21
SLAVE	82.70	85.50	86.80	87.20	82.70	86.80	86.80	84.30	85.10	82.30	84.80
GP-FCS	78.60	78.10	78.10	79.80	78.60	77.70	79.40	81.40	83.90	79.00	79.50
ЛДА	81.40	83.50	83.90	84.70	83.50	83.50	86.00	83.50	85.10	85.10	83.90
Д-НСНП	83.54	85.19	85.54	87.24	86.01	81.89	86.42	85.19	84.77	87.24	85.30
П-НСНП	86.83	85.19	85.54	86.42	86.01	86.83	86.83	85.19	85.19	86.42	86.04
Дані тестування											
С4.5	77.70	88.80	85.10	85.10	88.80	81.40	74.00	96.20	77.70	77.70	83.30
ШНМ	77.70	96.20	85.10	74.50	77.70	77.70	77.70	96.20	70.30	81.40	81.40
HF-GBML	77.78	96.30	85.19	70.37	92.59	77.78	77.78	93.10	70.37	81.48	82.27
SLAVE	74.00	96.20	81.40	77.70	81.40	70.30	74.00	92.50	77.70	77.70	80.30
GP-FCS	70.30	85.10	88.80	77.70	77.70	81.40	74.00	88.80	59.20	77.70	78.10
ЛДА	77.70	92.50	88.80	81.40	92.50	74.00	70.30	92.50	77.70	77.70	82.50
Д-НСНП	70.37	96.30	85.19	85.19	88.89	66.67	77.78	96.30	74.07	77.78	81.85
П-НСНП	77.78	96.30	85.19	85.19	88.89	81.48	77.78	96.30	77.78	77.78	84.44



ІНТЕГРАЦІЯ В ІНФОРМАЦІЙНУ СИСТЕМУ

На цьому етапі розроблена модель була інтегрована в інформаційну систему підтримки прийняття рішень, яка призначена для медичних установ. Система дозволяє лікарям вводити дані пацієнта, які можуть надходити як вручну, так і автоматично з медичних пристроїв (наприклад, ЕКГ-апаратів чи моніторів тиску). Дані обробляються на сервері, побудованому на сучасних технологіях, таких як Node.js, що забезпечує швидку й масштабовану обробку. Результати діагностики та аналізу зберігаються в базі даних MongoDB, що дозволяє швидкий доступ до інформації та історії пацієнтів.

Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Введення даних пацієнта

Ім'я пацієнта

Вік

Стать Жіноча Чоловіча

Тип болю в грудях

Артеріальний тиск

Рівень холестерину в крові

Результати ЕКГ 0 1 2

Рівень глюкози в крові натще (>120 мг/дл)

Максимальна частота серцевих скорочень

Стенокардія при навантаженні Ні Так

Депресія ST сегмента

Нахил ST сегмента під час фізичних навантажень

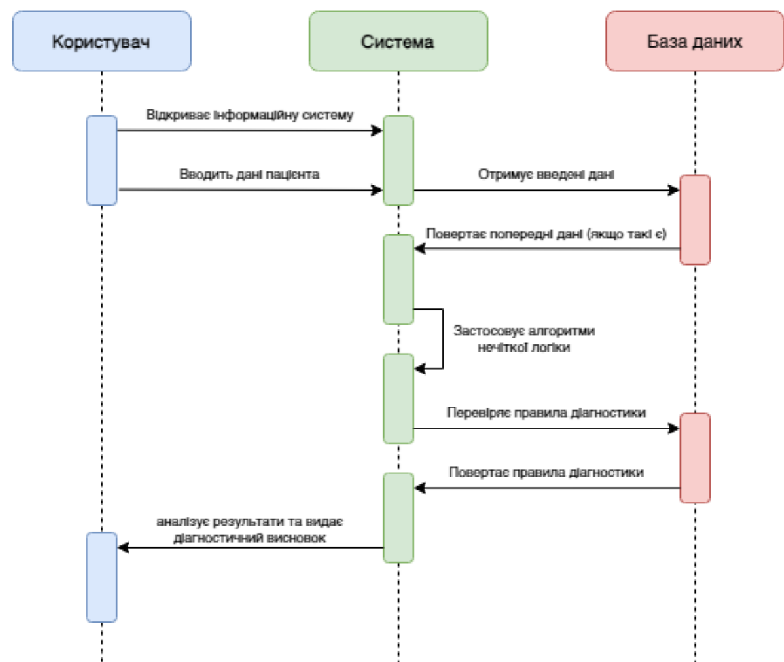
Кількість великих судин

Таласемія

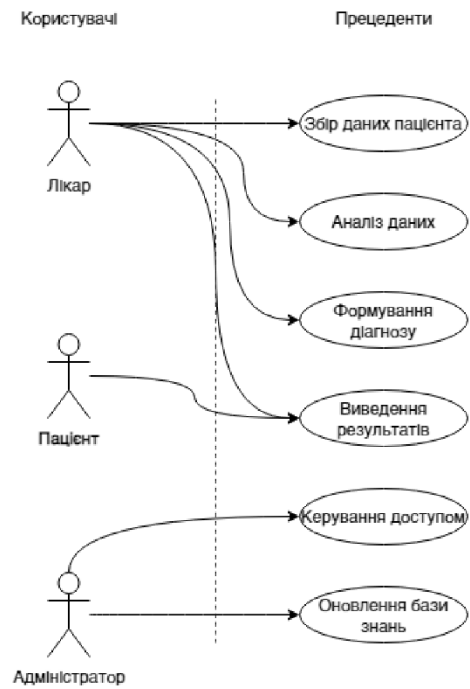
Діагностика ішемічної хвороби серця

Результати діагностики

ДІАГРАМА ПОСЛІДОВНОСТІ



USE—CASE ДІАГРАМА



ВИСНОВКИ

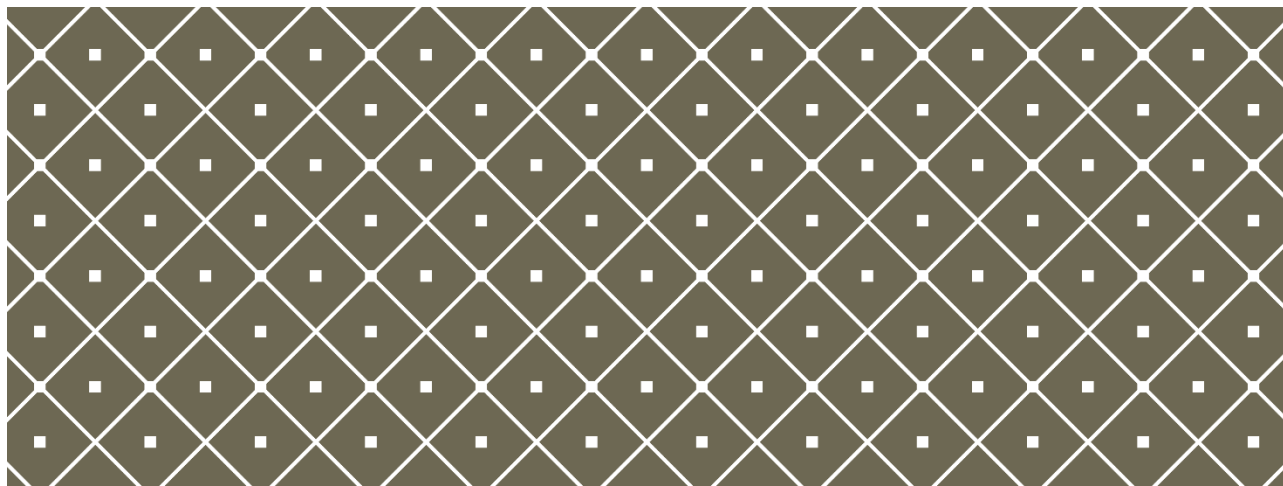
У дослідженні розроблено систему для діагностики ішемічної хвороби серця (ІХС), засновану на нечітких правилах, що використовує зрозумілі для лікаря терміни.

Система включає додаткові показники, зокрема ступінь достовірності правил, що дає можливість лікарю перевіряти їх обґрунтованість.

Для підвищення достовірності діагностики застосовано стратегію ансамблю класифікаторів (САК), що забезпечує залучення додаткових класифікаторів у разі низької визначеності.

Це сприяє підвищенню надійності діагностики та зниженню ризику помилок у процесі прийняття рішення.

0



ДЯКУЮ ЗА УВАГУ! |

ДОДАТОК В

РЕЗУЛЬТАТИ

Таблиця В.1 – Достовірність класифікації НСНП до (Д-НСНП) та після (П-НСНП) застосування стратегії ансамблевого класифікатора (САК) та низки еталонних алгоритмів класифікації з використанням методу 10-кратної перехресної перевірки

Методи	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	В середньому
Навчальні дані											
C4.5	86.80	85.10	86.40	86.40	86.00	86.80	86.80	85.10	87.20	86.40	86.30
ШНМ	87.60	85.50	86.80	87.20	86.00	87.20	87.60	85.50	87.60	87.20	86.80
HF-GBML	86.83	85.19	85.19	86.42	86.01	86.01	87.24	85.60	86.83	86.83	86.21
SLAVE	82.70	85.50	86.80	87.20	82.70	86.80	86.80	84.30	85.10	82.30	84.80
GP-FCS	78.60	78.10	78.10	79.80	78.60	77.70	79.40	81.40	83.90	79.00	79.50
ЛДА	81.40	83.50	83.90	84.70	83.50	83.50	86.00	83.50	85.10	85.10	83.90
Д-НСНП	83.54	85.19	85.54	87.24	86.01	81.89	86.42	85.19	84.77	87.24	85.30
П-НСНП	86.83	85.19	85.54	86.42	86.01	86.83	86.83	85.19	85.19	86.42	86.04
Дані тестування											
C4.5	77.70	88.80	85.10	85.10	88.80	81.40	74.00	96.20	77.70	77.70	83.30
ШНМ	77.70	96.20	85.10	74.50	77.70	77.70	77.70	96.20	70.30	81.40	81.40
HF-GBML	77.78	96.30	85.19	70.37	92.59	77.78	77.78	93.10	70.37	81.48	82.27
SLAVE	74.00	96.20	81.40	77.70	81.40	70.30	74.00	92.50	77.70	77.70	80.30
GP-FCS	70.30	85.10	88.80	77.70	77.70	81.40	74.00	88.80	59.20	77.70	78.10
ЛДА	77.70	92.50	88.80	81.40	92.50	74.00	70.30	92.50	77.70	77.70	82.50
Д-НСНП	70.37	96.30	85.19	85.19	88.89	66.67	77.78	96.30	74.07	77.78	81.85
П-НСНП	77.78	96.30	85.19	85.19	88.89	81.48	77.78	96.30	77.78	77.78	84.44

Таблиця В.2 – Значення ППК для НСНП до (Д-НСНП) та після (П-НСНП) застосування стратегії ансамблевого класифікатора (САК) та низки еталонних алгоритмів класифікації з використанням методу 10-кратної перехресної перевірки

Методи	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	В середньому
Навчальні дані											
C4.5	0.863	0.844	0.859	0.858	0.853	0.864	0.864	0.846	0.869	0.859	0.858
ШНМ	0.843	0.837	0.854	0.835	0.819	0.856	0.852	0.824	0.802	0.847	0.837
HF-GBML	0.863	0.844	0.849	0.858	0.853	0.860	0.868	0.849	0.864	0.863	0.857
SLAVE	0.819	0.848	0.862	0.866	0.815	0.836	0.865	0.834	0.844	0.811	0.840
GP-FCS	0.795	0.844	0.805	0.797	0.831	0.800	0.854	0.764	0.815	0.860	0.817
ЛДА	0.809	0.834	0.840	0.846	0.834	0.817	0.859	0.835	0.854	0.849	0.838
Д-НСНП	0.826	0.844	0.847	0.866	0.853	0.814	0.860	0.846	0.841	0.867	0.846
П-НСНП	0.863	0.844	0.847	0.858	0.853	0.864	0.865	0.846	0.844	0.859	0.854
Дані тестування											
C4.5	0.767	0.892	0.842	0.850	0.900	0.800	0.742	0.958	0.750	0.783	0.828
ШНМ	0.883	0.933	0.842	0.825	0.833	0.800	0.675	0.925	0.692	0.783	0.819
HF-GBML	0.767	0.967	0.842	0.692	0.933	0.767	0.775	0.917	0.675	0.817	0.815
SLAVE	0.717	0.967	0.800	0.767	0.808	0.675	0.742	0.917	0.750	0.783	0.793
GP-FCS	0.750	0.892	0.842	0.742	0.833	0.708	0.642	0.842	0.633	0.817	0.770
ЛДА	0.767	0.933	0.883	0.825	0.933	0.717	0.708	0.925	0.758	0.783	0.823
Д-НСНП	0.683	0.967	0.842	0.850	0.900	0.642	0.775	0.958	0.717	0.783	0.812
П-НСНП	0.767	0.967	0.842	0.850	0.900	0.800	0.775	0.958	0.750	0.783	0.839

Таблиця В.3 – Значення чутливості НСНП до (Д-НСНП) та після (П-НСНП) застосування стратегії ансамблевого класифікатора (САК) та низки еталонних алгоритмів класифікації з використанням методу 10-кратної перехресної перевірки

Методи	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	В середньому
Навчальні дані											
C4.5	0.815	0.778	0.815	0.806	0.787	0.824	0.824	0.796	0.843	0.815	0.810
ШНМ	0.833	0.815	0.833	0.833	0.815	0.852	0.852	0.833	0.833	0.806	0.831
HF-GBML	0.815	0.778	0.776	0.806	0.789	0.861	0.824	0.787	0.824	0.815	0.807
SLAVE	0.741	0.778	0.806	0.806	0.704	0.769	0.833	0.750	0.778	0.704	0.767
GP-FCS	0.694	0.778	0.676	0.750	0.833	0.704	0.870	0.676	0.741	0.787	0.751
ЛДА	0.759	0.824	0.843	0.833	0.824	0.796	0.852	0.833	0.870	0.824	0.826
Д-НСНП	0.741	0.778	0.785	0.806	0.787	0.769	0.824	0.796	0.778	0.815	0.788
П-НСНП	0.815	0.778	0.785	0.806	0.787	0.824	0.833	0.796	0.778	0.815	0.802
Дані тестування											
C4.5	0.667	0.917	0.750	0.833	1.000	0.667	0.750	0.917	0.500	0.833	0.783
ШНМ	0.833	1.000	0.750	0.917	1.000	0.667	0.750	0.917	0.583	0.833	0.825
HF-GBML	0.667	0.875	0.750	0.583	1.000	0.667	0.750	0.833	0.417	0.833	0.738
SLAVE	0.500	1.000	0.667	0.667	0.750	0.417	0.750	0.833	0.500	0.833	0.692
GP-FCS	0.500	0.917	0.750	0.750	1.000	0.417	0.750	0.750	0.333	0.833	0.700
ЛДА	0.667	1.000	0.833	0.917	1.000	0.500	0.750	0.917	0.583	0.833	0.800
Д-НСНП	0.500	1.000	0.750	0.833	1.000	0.417	0.750	0.917	0.500	0.833	0.750
П-НСНП	0.667	1.000	0.750	0.833	1.000	0.667	0.750	0.917	0.500	0.833	0.792

Таблиця В.4 – Значення специфічності НСНП до (Д-НСНП) та після (П-НСНП) застосування стратегії ансамблевого класифікатора (САК) та низки еталонних алгоритмів класифікатора з використанням методу 10-кратної перехресної перевірки

Методи	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	В середньому
Навчальні дані											
C4.5	0.910	0.911	0.904	0.911	0.919	0.904	0.904	0.896	0.896	0.904	0.906
ШНМ	0.852	0.859	0.874	0.837	0.822	0.859	0.852	0.815	0.770	0.889	0.843
HF-GBML	0.911	0.911	0.933	0.911	0.907	0.859	0.911	0.911	0.904	0.911	0.907
SLAVE	0.896	0.919	0.919	0.926	0.926	0.904	0.896	0.919	0.911	0.919	0.913
GP-FCS	0.896	0.911	0.933	0.844	0.830	0.896	0.837	0.852	0.889	0.933	0.882
ЛДА	0.859	0.844	0.837	0.859	0.844	0.837	0.867	0.837	0.837	0.874	0.850
Д-НСНП	0.911	0.911	0.911	0.926	0.919	0.859	0.896	0.896	0.904	0.919	0.905
П-НСНП	0.911	0.911	0.911	0.911	0.919	0.904	0.896	0.896	0.911	0.904	0.907
Дані тестування											
C4.5	0.867	0.867	0.933	0.867	0.800	0.933	0.733	1.000	1.000	0.733	0.873
ШНМ	0.933	0.867	0.933	0.733	0.667	0.933	0.600	0.933	0.800	0.733	0.813
HF-GBML	0.867	0.917	0.933	0.800	0.867	0.867	0.800	1.000	0.933	0.800	0.878
SLAVE	0.933	0.933	0.933	0.867	0.867	0.933	0.733	1.000	1.000	0.733	0.893
GP-FCS	1.000	0.867	0.933	0.733	0.667	1.000	0.533	0.933	0.933	0.800	0.840
ЛДА	0.867	0.867	0.933	0.733	0.867	0.933	0.667	0.933	0.933	0.733	0.847
Д-НСНП	0.867	0.933	0.933	0.867	0.800	0.867	0.800	1.000	0.933	0.733	0.873
П-НСНП	0.867	0.933	0.933	0.867	0.800	0.933	0.800	1.000	1.000	0.733	0.887

Anti-Plagiarism v-15.257

Максимальне співпадіння з одним документом 1.0%

Словники перевірки: en_US, ru_RU, ua_UA. Помилоч в документах: 5%

ID: 156926 Назва: МКР Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу Додано в БД: 2024-12-09 Автора: Владислав КУЛЬБАЧНИЙ Керівники: Андрій НІЧЕПОРУК Консультанти: Опоненти:	Документ		Сумарний збіг по Базі Даних	
	Символи	Лексеми	Символи	Лексеми
	115038	907	3280 (3%)	56 (6%)

Джерело плагіату

ID	Опис	Наявність плагіату в документі	
		Символи	Лексеми

Протокол аналізу звіту подібності експертом

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

Автор: Кульбачний В.

Співавтор:

Назва: Кульбачний_Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Експерт: Нічепорук А.О.

Підрозділ: Кафедра комп'ютерної інженерії та інформаційних систем

Коефіцієнт подібності 1: 1.3%

Коефіцієнт подібності 2: 0.3%

Мікропробіли: 58

Заміна букв: 0

Інтервали: 0

Білі знаки: 1

Дата створення звіту: 2024-12-09 17:44:19.0

Після аналізу Звіту подібності констатую наступне:

Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.

Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.

Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедур. Таким чином робота не приймається.

Обґрунтування:

2024-12-09

Дата



Доцент Андрій Нічепорук

експерт

РЕЦЕНЗІЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

Дипломник: Кульбачний Владислав Васильович

Тема: Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Спеціальність: 126 «Інформаційні системи та технології»

Обсяг дипломної роботи:

Кількість листів креслень —; кількість сторінок записки 74

1. Короткий зміст роботи та прийнятих рішень У роботі запропоновано інформаційну систему підтримки прийняття рішення для визначення медичного діагнозу

2. Висновок про відповідність роботи дипломному завданню _____

Кваліфікаційна _____ робота _____ відповідає _____ виданому завданню _____

3. Характеристика виконання кожного розділу, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У першому розділі проведено аналіз відомих рішень для оптимізації енергоспоживання у Розумному будинку. У першому розділі представлено аналіз сучасних підходів до побудови інформаційних систем підтримки прийняття рішень у медичній діагностиці. Другий розділ присвячений вибору методів машинного навчання для діагностики ІХС. У третьому розділі описано архітектуру та реалізацію прототипу інформаційної системи. У четвертому розділі наведено результати експериментальних досліджень і оцінку ефективності системи.

4. Позитивні сторони роботи: Запропоновано інформаційну систему підтримки прийняття рішень для діагностики ішемічної хвороби серця, розроблено модель процесу діагностики із використанням сучасних методів машинного навчання, описано структуру і компоненти інформаційної системи, а також створено прототип програмного забезпечення, що реалізує запропоновані підходи.

5. Негативні сторони роботи: Основним недоліком запропонованої інформаційної системи є її складність у впровадженні через потребу в значній кількості даних для навчання моделей машинного навчання. Крім того, низька інтерпретованість деяких алгоритмів може викликати труднощі у використанні системи лікарями без відповідної підготовки, що може вплинути на рівень довіри до її рішень.

6. Оцінка графічного оформлення та пояснювальної записки роботи: _____

7. Відгук про роботу в цілому: В загальному робота виконана на достатньому рівні.

8. Інші зауваження: _____

9. Оцінка дипломної роботи:

Розглянувши позитивні та негативні сторони представленої дипломної роботи вважаю, що робота заслуговує оцінки «добре» 4,0 (С)

Рецензент (прізвище, ім'я, по батькові, посада, місце роботи)
Радюк Л.М., Ph.D., ст. викладач каф. КН

«11» 12 2024р.



В.о. завідувачу кафедри КІС
ДФ, доц. Павловій О.О.

Кульбачний Владислав Васильович

ГІБ здобувача вищої освіти

ФІТ, 2 курсу, групи ІСТМ-23-1

ЗАЯВА

З правилами чинного Положення про систему забезпечення академічної доброчесності у Хмельницькому національному університеті, згідно з яким виявлення академічного плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту і застосування заходів дисциплінарної та академічної відповідальності, ознайомлений (а). Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на наявність академічного плагіату оповіщений (а) та надаю свою згоду на обробку й збереження університетом моєї роботи в інституційному репозитарії Хмельницького національного університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-обчислювального комплексу StrikePlagiarism та/або програмно-технічного засобу Anti-Plagiarism) і використання роботи для виявлення академічного плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення текстових збігів в роботах.

Робота надається для перевірки в електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

06 грудня 2024 року



РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ
КАФЕДРИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Підтверджуємо ознайомлення з результатами звіту/звітів перевірки роботи, продукованими програмно-технічним засобом (ами), на наявність текстових збігів:

Назва кваліфікаційної роботи: Інформаційна система підтримки прийняття рішення для визначення медичного діагнозу

Автор: Кульбачний Владислав Васильович

Освітня програма: освітньо-професійна

Спеціальність: 126 – Інформаційні системи та технології

Науковий керівник: Нічепорук А.О., к.т.н, доцент

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є академічним плагіатом. Робота приймається до захисту.	відповідає
2	Виявлені запозичення не є академічним плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота приймається до захисту, але має бути відкоригована.	
3	Виявлені запозичення не є академічним плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту після того як буде відкоригована та доопрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	

Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

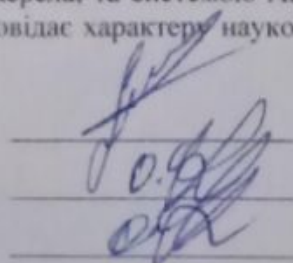
- 1) усі запозичення фрагментарні, або мають належним чином оформленні посилання;
- 2) окремі виявлені збіги є загальноновживаними фразами або виразами, про що свідчить посилання системи на збіг з джерелами на один фрагмент речення;

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів StrikePlagiarism, складає 1,27% і адресується до 15 першоджерела; та системою Anti-Plagiarism складає 1%, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Керівник роботи

Гарант ОП

В.о. завідувача кафедри КІС



А.О. Нічепорук

О. О. Павлова

О. О. Павлова