

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

(повне найменування вищого навчального закладу)

Факультет комп'ютерно-інформаційних систем і програмної інженерії

(назва факультету)

Кафедра комп'ютерних систем та мереж

(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістра

(освітній ступінь)

на тему:

**Методи та програмно-апаратні засоби опрацювання
даних в системах моніторингу рівня цукру в крові**

Виконав: студент (ка) 6 курсу, групи СІМ-61

спеціальності

123 «Комп'ютерна інженерія»

(шифр і назва спеціальності)

Керівник

(підпис)

Баран С.О.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Луцків А.М.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Луцик Н.С.

(прізвище та ініціали)

Рецензент

(підпис)

Осухівська Г.М.

(прізвище та ініціали)

Цуприк Г.Б.

(прізвище та ініціали)

Тернопіль
2022

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

(повне найменування вищого навчального закладу)

Факультет комп'ютерно-інформаційних систем і програмної інженерії

Кафедра комп'ютерних систем та мереж

ЗАТВЕРДЖУЮ

Завідувач кафедри Осухівська Г.М.

«_____»

2022 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня магістр
(назва освітнього ступеня)

за спеціальністю 123 «Комп'ютерна інженерія»
(шифр і назва спеціальності)

студенту Барану Сергію Олександровичу
(прізвище, ім'я, по-батькові)

1. Тема проекту (роботи) Методи та програмно-апаратні засоби опрацювання даних в системах моніторингу рівня цукру в крові

Керівник проекту (роботи) Луцків Андрій Мирославович, к.т.н., доц.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «06» грудня 2022 року №4/7-986

2. Термін подання студентом завершеної роботи _____

3. Вихідні дані до роботи Методи визначення діабету, принципи та особливості систем вимірювання рівня глюкози в крові, особливості накопичення медичних даних і прогнозування

4. Зміст роботи (перелік питань, які потрібно розробити)

Вступ. 1. Аналіз сучасного стану досліджень у сфері виявлення, аналізу і прогнозування розвитку цукрового діабету 2. Побудова архітектури програмно-апаратних засобів опрацювання даних в системах моніторингу рівня цукру в крові 3. Програмна модель прогнозування розвитку цукрового діабету 4. Охорона праці та безпека в надзвичайних ситуаціях. Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Актуальність і мета дослідження. 2. Задачі дослідження, об'єкт і предмет, наукова новизна і практична цінність дослідження. 3. Статистика захворюваності цукровим діабетом. 4. Структура та алгоритм роботи вбудованого неінвазивного глюкометра. 5. Архітектура інформаційної системи управління медичними даними 6. Модель архітектури розподіленої системи моніторингу і прогнозування рівня цукру в крові 7. Результати оцінювання точності прогнозування розвитку цукрового діабету та фактори впливу на його розвиток. 8. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
<i>Охорона праці та безпека в надзвичайних ситуаціях</i>	<i>Осухівська Г.М.</i>		

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	<i>Аналіз сучасного стану досліджень у сфері виявлення, аналізу і прогнозування розвитку цукрового діабету</i>		<i>виконано</i>
2.	<i>Побудова архітектури програмно-апаратних засобів опрацювання даних у системах моніторингу рівня цукру в крові</i>		<i>виконано</i>
3.	<i>Програмна модель прогнозування розвитку цукрового діабету</i>		<i>виконано</i>
4.	<i>Охорона праці та безпека в надзвичайних ситуаціях</i>		<i>виконано</i>
5.	<i>Оформлення пояснювальної записки</i>		<i>виконано</i>
6.	<i>Оформлення графічного матеріалу</i>		<i>виконано</i>
7.	<i>Попередній захист кваліфікаційної роботи магістра</i>		<i>виконано</i>
8.	<i>Захист кваліфікаційної роботи магістра</i>		

Студент

(підпис)*Баран С.О.*_____
(прізвище та ініціали)

Керівник проекту (роботи)

(підпис)*Луцків А.М.*_____
(прізвище та ініціали)

АНОТАЦІЯ

Тема кваліфікаційної роботи: “ Методи та програмно-апаратні засоби опрацювання даних в системах моніторингу рівня цукру в крові ” // Кваліфікаційна робота // Баран Сергій Олександрович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп’ютерно-інформаційних систем та програмної інженерії, група СІМ-61 // Тернопіль, 2022 // с. – 90, рис. – 34, табл. – 5, аркушів А1 – 8, додат. – 1, бібліогр. – 26.

Ключові слова: метод, засіб, опрацювання, моніторинг, діабет.

Мета кваліфікаційної роботи полягає у дослідженні методів і програмно-апаратних засобів опрацювання даних в системах моніторингу рівня цукру в крові.

У кваліфікаційній роботі проаналізовано існуючі підходи, методи та інструменти для визначення рівня цукру в крові людини і встановлено, що найбільш поширеними для побутового використання є пристрої інвазивного та неінвазивного вимірювання, а системи моніторингу в Україні відсутні або недосконалі, що не дає змоги забезпечити якість обслуговування пацієнтів та виконувати прогнозування розвитку цього захворювання.

Розроблено алгоритми функціонування глюкометра та глобальної інформаційної системи управління медичними даними, які в комплексі становлять систему цілодобового моніторингу та керування рівнем глюкози в крові пацієнта.

Побудовано та математично представлено концептуальну модель розподіленої архітектури системи збору та опрацювання даних для моніторингу рівня цукру в крові, що включає в себе сукупність локальних і центрального вузла управління і дає змогу обмінюватися повідомленнями та виконувати прогнозування розвитку хвороби. Реалізовано алгоритми прогнозування (класифікації) розвитку цукрового діабету на основі підходів дерев прийняття рішень, випадкового лісу, логістичної регресії та XGBoost.

ABSTRACT

The theme of the thesis: " Methods, hardware and software for processing data in blood sugar monitoring systems" /Master thesis / Baran Serhii Oleksandrovysh / Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and software engineering, group CIm -61 // Ternopil, 2022// p. - 90, fig. – 34, table. – 5, Posters A1 – 8, Add – 1, Ref. – 26.

Keywords: method, tool, processing, monitoring, diabet.

The purpose of the qualification work is to research methods and hardware and software for processing data in blood sugar monitoring systems.

In the qualification work, the existing approaches, methods and tools for determining the level of sugar in human blood were analyzed and it was established that the most common for household use are devices for invasive and non-invasive measurement, and monitoring systems in Ukraine are absent or imperfect, which does not make it possible to ensure the quality of patient care and predict the development of this disease.

Algorithms for the operation of the glucometer and the global information system for managing medical data have been developed, which together form a system of round-the-clock monitoring and control of the patient's blood glucose level.

A conceptual model of the distributed architecture of the data collection and processing system for blood sugar monitoring, which includes a set of local and central control nodes and enables the exchange of messages and forecasting the development of the disease, is built and mathematically presented. Algorithms for predicting (classification) diabetes development based on decision tree, random forest, logistic regression and XGBoost approaches have been implemented.

ЗМІСТ

ВСТУП	8
РОЗДІЛ 1 АНАЛІЗ СУЧАСНОГО СТАНУ ДОСЛІДЖЕНЬ У СФЕРІ ВИЯВЛЕННЯ, АНАЛІЗУ І ПРОГНОЗУВАННЯ РОЗВИТКУ ЦУКРОВОГО ДІАБЕТУ	12
1.1. Аналіз статистики і факторів впливу на розвиток цукрового діабету	12
1.2. Аналіз існуючих типів приладів для вимірювання рівня цукру в крові	16
1.3. Висновки до розділу	24
РОЗДІЛ 2 ПОБУДОВА АРХІТЕКТУРИ ПРОГРАМНО-АПАРАТНИХ ЗАСОБІВ ОПРАЦЮВАННЯ ДАНИХ В СИСТЕМАХ МОНІТОРИНГУ РІВНЯ ЦУКРУ В КРОВІ	26
2.1. Структура та алгоритми роботи пристроїв вимірювання глюкози в крові.....	26
2.1.1. Система вбудованого глюкометра	26
2.1.2. Система збору та опрацювання показників рівня глюкози в крові	33
2.2. Можливі рішення щодо забезпечення розподіленості системи збору та опрацювання медичних даних	37
2.3. Побудова моделі розподіленої архітектури системи збору та опрацювання даних у системах моніторингу рівня цукру в крові.....	43
2.4. Висновки до розділу	47
РОЗДІЛ 3 ПРОГРАМНА МОДЕЛЬ ПРОГНОЗУВАННЯ РОЗВИТКУ ЦУКРОВОГО ДІАБЕТУ	49
3.1. Аналіз даних для прогнозування розвитку цукрового діабету	49
3.2. Виявлення кореляції між факторами впливу на розвиток діабету	56
3.3. Препроцесинг даних при прогнозуванні розвитку цукрового діабету.....	60
3.4. Реалізація алгоритмів прогнозування розвитку захворюваності цукровим діабетом.....	63
3.5. Висновки до розділу	73

РОЗДІЛ 4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ	75
4.1. Охорона праці	75
4.2. Особливості роботи та розлади здоров'я користувачів комп'ютерів, що формується під впливом роботи за комп'ютером.	78
ВИСНОВКИ.....	82
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	84
ДОДАТОК А ТЕЗИ КОНФЕРЕНЦІЙ	87

ВСТУП

Актуальність теми. Інформаційні технології сьогодні є рушієм розвитку та еволюції у багатьох сферах людської діяльності, починаючи від систем загального використання закінчуючи сферою критичних технологій. Від якості та надійності застосування програмно-апаратних комплексів залежить ефективність процесів, які вони дозволяють автоматизувати. До систем загального використання інформаційних технологій належать системи масової експлуатації, зокрема, програмні системи, які можна вільно завантажити з маркетплейсів і встановити локально як на мобільні пристрої, так і на комп'ютери користувачів. До критичних систем відносяться системи, які мають безпосередній вплив на життя і здоров'я людей, навколишнього середовища та ін. Особливо актуальним є впровадження інформаційних технологій у сфері медицини. Вони дають змогу аналізувати, проводити моніторинг та прогнозувати розвиток різноманітних хвороб, підвищувати якість життя пацієнтів, а також виступати в якості систем підтримки прийняття рішень для лікарів при встановленні діагнозу.

Сучасний етап розвитку людства супроводжується поширенням хвороб, які мають тенденції до зростання. Масового поширення набувають цукровий діабет, серцево-судинні захворювання, інсульты та інфаркти, хвороба Альцгеймера та розсіяний склероз, ракові захворювання, які в основному спричинені генетикою та особливостями способу життя, географічного проживання людей та ін.

Аналізуючи поширення цукрового діабету спостерігають тенденції щодо зростання кількості хворих до масштабів пандемії. Так, згідно з прогнозами міжнародних організацій, спостерігається семикратне зростання захворюваності у 2045 році порівняно з 2000 роком, що виражаються в абсолютних показниках як 700 млн. осіб щодо 145 млн. відповідно.

Зважаючи на такі невтішні показники, актуальною задачею на сьогодні є побудова комплексної системи визначення і прогнозування розвитку цукрового діабету.

Моніторинг рівня глюкози у крові є визначальною частиною життя для осіб хворих цукровим діабетом, тому важливо, в першу чергу, забезпечити зручність і точність спостереження за кількістю глюкози, автоматизовану передачу даних до центрального сховища і прогнозування його розвитку.

Тому актуальною задачею на сьогодні є побудова системи, до складу якої входить апаратне і програмне забезпечення, яке дозволить забезпечити збір і накопичення інформації, управління та прогнозування розвитку цукрового діабету.

Мета кваліфікаційної роботи полягає у дослідженні методів і програмно-апаратних засобів опрацювання даних в системах моніторингу рівня цукру в крові.

Об'єктом дослідження є процеси збору і накопичення даних, а також прогнозування розвитку цукрового діабету.

Предметом дослідження є методи і засоби накопичення даних та прогнозування розвитку цукрового діабету.

Для досягнення цієї мети у кваліфікаційній роботі магістра поставлено і наступні **задачі**:

- провести аналіз наукових публікацій щодо факторів впливу на розвиток цукрового діабету;
- дослідити існуючі програмно-апаратні засоби та інші рішення для виявлення та регулювання рівня глюкози в крові людини, накопичення та опрацювання таких даних;
- запропонувати архітектуру та можливі шляхи реалізації програмно-апаратної системи збору, накопичення та прогнозування рівня глюкози в крові;
- побудувати моделі та алгоритми прогнозування розвитку цукрового діабету на основі існуючих відкритих даних;
- імплементувати програмне забезпечення для прогнозування розвитку цукрового діабету.

Методи дослідження: При вирішенні задач кваліфікаційної роботи застосовувались такі методи і засоби: аналіз та узагальнення – при проведенні аналізу статистичних даних захворюваністю на цукровий діабет та виборі шляхів імплементатії системи моніторингу рівня цукру в крові; теорії множин та методів

машинного навчання – при формалізації та побудові концептуальної моделі розподіленої архітектури системи моніторингу та при прогнозуванні розвитку цукрового діабету; проектування та програмування – при реалізації програмної моделі прогнозування розвитку цукрового діабету; експеримент та вимірювання – при оцінюванні точності результатів прогнозування та виявленні факторів найбільшого впливу на розвитку цукрового діабету.

Наукова новизна отриманих результатів. Наукова новизна результатів дослідження полягає в наступному:

– уперше запропоновано алгоритми функціонування глюкометра та глобальної інформаційної системи управління медичними даними, які в комплексі становлять систему цілодобового моніторингу та керування рівнем глюкози в крові пацієнта, що дає змогу забезпечувати збір, опрацювання та прогнозування появи чи розвитку цукрового діабету;

– уперше побудовано та математично представлено концептуальну модель розподіленої архітектури системи збору та опрацювання даних для моніторингу рівня цукру в крові, що включає в себе сукупність локальних і центрального вузла управління і дає змогу обмінюватися повідомленнями та виконувати прогнозування розвитку хвороби.

Практичне значення одержаних результатів. Впровадження запропонованого рішення щодо використання програмного та апаратного забезпечення при реалізації систем моніторингу рівня цукру в крові дає змогу забезпечити цілодобове спостереження та прогнозування потенційної можливості появи і розвитку цукрового діабету.

Публікації. Результати кваліфікаційної роботи апробовані на X науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2022 року) як тези конференцій.

1. Луцків А.М., Баран С.О. Технології неінвазивного вимірювання рівня глюкози в крові. Матеріали X науково-технічної конференції Тернопільського

національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2022 року). Тернопіль: ТНТУ. 2022. С. 63.

2. Луцків А.М., Баран С.О. Алгоритми машинного навчання для прогнозування рівня глюкози в крові. Матеріали X науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2022 року). Тернопіль: ТНТУ. 2022. С. 64.

Структура роботи. Кваліфікаційна робота містить розрахунково-пояснювальну записку та графічний матеріал. До складу записки входить вступу, 4 розділи, загальні висновки, список використаних джерел і додатки. Обсяг роботи: розрахунково-пояснювальної записки – 95 арк. формату А4, графічна частина – 8 аркушів формату А1.

РОЗДІЛ 1

АНАЛІЗ СУЧАСНОГО СТАНУ ДОСЛІДЖЕНЬ У СФЕРІ ВИЯВЛЕННЯ,
АНАЛІЗУ І ПРОГНОЗУВАННЯ РОЗВИТКУ ЦУКРОВОГО ДІАБЕТУ

1.1. Аналіз статистики і факторів впливу на розвиток цукрового діабету

Цукровий діабет представляє собою хронічне метаболічне захворювання, що характеризується підвищеним рівнем глюкози в крові внаслідок абсолютного або відносного дефіциту інсуліну [1]. За оцінками Міжнародної федерації діабету, станом на 2022 рік 537 мільйонів дорослих людей у віці 20–79 років зараз живуть з діабетом у всьому світі. Очікується, що до 2030 року їх кількість зросте до 643 мільйонів [2].

Глобальність проблеми підвищеного рівня захворюваності цукровим діабетом сформувало складну медичну і соціальну проблему, що набуває ознак і характеру пандемії (рис. 1.1).



Рис. 1.1. Діабет як світова пандемія

Враховуючи глобальну тенденцію до виявлення і поширення цукрового діабету, цей тренд спостерігається і в Україні. Зокрема, за півтори десятка років у нас спостерігають зростання поширеності цієї хвороби більш ніж на п'ятдесят відсотків, а кількість випадків захворюваності зростала більш, ніж на 80%, що показано на рис. 1.2.



Рис. 1.2. Тренд захворюваності на цукровий діабет в Україні

Лікування діабету спрямоване на те, щоб допомогти людям з цією хворобою досягти майже нормального рівня глікемії для зменшення ризику довгострокових (наприклад, судинних) ускладнень, уникаючи при цьому гострих метаболічних ризиків і зберігаючи найкращу якість життя.

Розрізняють різні види і стадії хвороби. Для прикладу, діабет першого типу зумовлений автоімунною реакцією при якій організм людини пошкоджує клітини які, відповідають за інсулін. Це впливає на його недостатність та порушення функцій організму. Імовірними причинами розвитку цукрового діабету є специфічна структура на генному рівні, в якості тригерів, які стимулюють розвиток діабету є фактори навколишнього середовища, вірусні інфекції та порушення імунної системи (рис. 1.3).



Рис. 1.3. Фактори, які впливають на розвиток цукрового діабету

Ключовим фактором для досягнення хорошого рівня глікемії є самоконтроль хворими свого стану. Особи з цукровим діабетом повинні:

- контролювати споживання вуглеводів, шляхом вибору їжі, адаптації харчової поведінки щодо глікемічного навантаження;
- дотримуватися принципів здорового харчування;
- керувати кількістю глюкози у крові за допомогою глюкозо-знижуючих препаратів;
- моніторити рівень цукру за допомогою традиційного аналізу крові або комп'ютерних систем з відповідними сенсорами;
- забезпечувати фізичне навантаження для оптимізації глікемії та контролю маси тіла;
- організувати діяльність відповідно до поточних рівнів глікемії та вимог до лікування, які рекомендовано рекомендаціями лікарів.

Якщо використовується інсулін швидкої дії (для покриття підвищеного рівня глюкози після прийняття їжі), оцінка вуглеводного навантаження, коригування дози інсуліну та коригування підвищених рівнів глюкози є додатковою необхідною практикою щоденного самоконтролю при діабеті.

Стійка або рецидивуюча гіперглікемія підвищує ризик розвитку серйозних віддалених ускладнень діабету, зокрема діабетичної ретинопатії, нейропатії, нефропатії з синдромом стопи.

Крім того, субоптимальний контроль глікемії пов'язаний із підвищеним ризиком гострих метаболічних ускладнень, таких як важка гіпоглікемія або важка гіперглікемія з ризиком кетоацидозу або гіперосмолярної коми [8].

Таким чином, прийняття та підтримка поведінки функціонального самоконтролю для досягнення хорошого глікемічного результату є вирішальним для підтримки доброго здоров'я та запобігання ускладненням і захворюваності [9]. Проте деякі дані підтверджують, що практики самоконтролю та загальна продуктивність людей із діабетом часто можна покращити, зокрема, це може бути особливо корисним для людей із супутніми психічними захворюваннями, такими як депресія та характерний для діабету дистрес [11].

Оскільки, самоконтроль є вирішальним фактором, що визначає перебіг цукрового діабету, відображення/моніторинг відповідної поведінки окремих осіб для виявлення областей потенційного покращення та надання відповідної підтримки може бути корисним у рутинній клінічній практиці.

Оцінка самоконтролю при захворюванні діабетом може представляти особливий інтерес для людей із стійкими субоптимальними результатами діабету, де необхідно виявити можливі проблеми та загрози. Крім того, вимірювання може знадобитися як частина дослідження, де аналізуються сприятливі фактори для оптимізації лікування діабету, включаючи психічні фактори [12] або вплив втручань (наприклад, навчання самоконтролю діабету) необхідно оцінити. Тому одним з важливих факторів при самоконтролі є впровадження простих та ефективних пристроїв для вимірювання рівня глюкози в крові.

Результати системного аналізу доступних інструментів вимірювання при самоконтролі діабету підтверджують, що на сьогодні створено багато різних засобів, однак, їх більшість застосовувалася в обмеженій кількості досліджень, а тестування вимірювальних властивостей часто були обмеженими, з невеликою

кількістю шкал оцінювання, які відповідали строгим критеріям, згідно з висновками експертів.

Ці проблеми обмежують застосування існуючих вимірювальних інструментів. У 2013 році було запроваджено опитувальник самоконтролю діабету, щоб забезпечити багатофакторність для оцінювання поведінки діабету, що відіграє важливу роль з метою моніторингу глікемії в основних типах діабету.

У прямих порівняннях DSMQ пояснював значно більшу варіацію глікемії, ніж встановлена стандартна шкала самообслуговування [13]. Відтоді її перекладено різними мовами та використано в багатьох дослідженнях, підтверджуючи її потенційну цінність для досліджень і практики. У нещодавньому системному огляді DSMQ було визначено одну з трьох шкал самоконтролю діабету, які відповідають рекомендаціям COSMIN щодо засобів вимірювання, які можна рекомендувати для використання, і результатам, отриманих за їх допомогою можна цілком довіряти [14].

Однак технологічні інновації, такі як постійний моніторинг рівня глюкози та автоматична доставка інсуліну, змінили терміни та шляхи в лікуванні діабету. Крім того, інструмент, який відповідає вимогам, встановлених організацією [7], має краще охопити деякі конкретні аспекти самоуправління.

Згідно [12], в Україні наразі немає можливості щодо побудови комплексної системи аналізу і прогнозування тренду захворюваності на цукровий діабет, оскільки відсутні організаційні механізми і технічні засоби для формування та опрацювання статистики щодо розвитку діабету та смертності від цього захворювання.

1.2. Аналіз існуючих типів приладів для вимірювання рівня цукру в крові

Перевірка рівня глюкози в крові (цукру в крові) – єдиний точний спосіб дізнатися його рівень та визначити чи він занадто високий чи низький. Звичайна людина не може точно визначити рівень цукру у своїй крові лише за своїм самопочуттям. Його можна перевірити за допомогою спеціалізованого обладнання.

Глюкометр – прилад, за допомогою якого береться крапля крові з пальця людини, яку надягають на спеціальну паличку. Потім прилад використовується для визначення рівня глюкози в крові.

Флеш-моніторинг рівня глюкози – ця система використовує сенсорну технологію для перевірки рівня глюкози без необхідності колоти палець. Датчик, який носять на тильній стороні руки, підключається до пристрою для зчитування або додатку для мобільного телефону, який зберігає і моніторить рівень глюкози вдень і вночі.

Пристрій безперервного моніторингу рівня глюкози (CGM) – під шкіру встановлюється датчик, який вимірює рівень глюкози 24 години на добу. CGM корисний для людей, які мають постійні проблеми з контролем рівня цукру в крові. Вартість, включаючи витратні матеріали (датчики), становить близько 5000 доларів на рік.

Для прикладу, уряд Австралії надає доступ до субсидованих безперервних і миттєвих продуктів для моніторингу рівня глюкози через Національну схему служб діабету (NDSS). Продукти, які субсидуються NDSS, такі як голки, шприци, смужки для моніторингу рівня глюкози в крові, смужки для моніторингу сечі та витратні матеріали для інсулінової помпи, можна придбати в точках доступу NDSS, як правило, в аптеці.

Обладнання для тестування кетонів. Не всім хворим на цукровий діабет потрібно перевіряти кетони. В основному це пацієнти, які використовують інсулін. Якщо потрібно провести тест на кетони, можна скористатися наступними засобами:

Тестові смужки або палички – їх занурюють у сечу та вони змінюють колір, щоб показати кількість кетонів в організмі людини.

Глюкометри –перевіряють кетони у крові та працюють подібно до вимірювання рівня глюкози в крові.

Проблеми з обладнанням для тестування та моніторингу. Глюкометри можуть перестати працювати належним чином, якщо:

- лічильник стає старим;

- лічильник перегрівається, вологий або брудний;
- батареї потребують заміни або підзарядки;
- тест-смужки застаріли;
- код калібрування неправильний;
- використання несумісної з глюкометром смужки;
- на смужці недостатньо крові або смужку неправильно вставлено в глюкометр;
- на руках було щось солодке перед тестом.

Щоб уникнути проблем із глюкометром преш за все потрібно ознайомитися з інструкцією виробника щодо правильного догляду та використання приладу. Перед тестуванням необхідно помити руки з милом, потім сполоснути і висушити їх, оскільки бруд може спричинити неточні покази.

Безперервний моніторинг рівня глюкози (CGM). При використанні CGM датчик потрібно знімати та встановлювати новий на іншу частину тіла раз на тиждень. Варто регулярно перевіряти рівень цукру в крові за допомогою аналізу крові з пальця, щоб переконатися, що CGM вимірюється належним чином (тобто переконатися, що показники збігаються).

Флеш-моніторинг глюкози. Важливо переконатися, що шкіра чиста та суха, перш ніж застосовувати флеш-монітор глюкози, інакше він може не працювати належним чином.

Батареї для діабетичного обладнання, зазвичай, можна придбати в будь-якому магазині, який продає електроніку. Однак потрібно перевіряти тип необхідної батареї та способи її заміни.

Діабет можна діагностувати, якщо рівень глюкози в крові натще становить 126 мг/дл або вище. Нормальний результат тесту на глюкозу натще нижче 100 мг/дл. Однією з головних цілей лікування діабету є підтримання рівня глюкози в крові в межах заданого цільового діапазону.

Більше 400 мільйонів людей у всьому світі живуть з діабетом, і вони все ще страждають від того, що кілька разів на день колють пальці, щоб перевірити рівень глюкози в крові. Різні методи, альтернативні методу проколу пальця, були широко

вивчені для визначення рівня глюкози в крові, зокрема це стосується ферментного або оптичного сенсорів глюкози.

Тим не менш, вони все ще мають проблеми з точки зору довговічності, портативності та точності. У дослідженні [14] дослідницька група запровадила напівпостійний і безперервний контроль рівня цукру в крові з низькими витратами на обслуговування без болю, спричиненого забором крові, що дозволяє пацієнтам забезпечити якість життя завдяки належному лікуванню та контролю діабету.

Очікується, що це збільшить використання CGMS, яке зараз становить лише 5%. Дослідницька група також провела як внутрішньовенний тест на толерантність до глюкози (IVGTT), так і оральний тест на толерантність до глюкози (OGTT) з датчиком, імплантованим свиням у контрольованому середовищі.

За словами дослідницької групи [14], результати початкового експерименту з підтвердження концепції «in vivo» показали багатообіцяючу кореляцію між рівнем цукру в крові і частотною характеристикою датчика. Сенсор демонструє здатність відстежувати тенденцію рівня цукру в крові і для фактичної імплантації датчика потрібно розглянути біосумісну упаковку та реакцію сторонніх тіл для тривалого застосування. Крім того, вдосконалена система інтерфейсу сенсора знаходиться в стадії розробки.

Потрібно відмітити, що пристрої, які призначені для постійного моніторингу рівня цукру в крові на сьогодні практично не використовуються в Україні і лише 5% такого типу обладнання застосовується у світі. Це пов'язано з їх вартістю і зручністю для пацієнта. Проте, дуже поширеними є два види глюкометрів: інвазивні та неінвазивні.

Характерною особливістю пристроїв неінвазивного вимірювання рівня глюкози є те, що вони не пошкоджують покриви шкіри і вимірювання можна проводити частіше, ніж з використанням традиційних глюкометрів. Однак недоліком таких засобів є те, що точність їхнього вимірювання може мати значні похибки у випадку порушення кровопостачання, наявності загрубілої шкіри чи мозолів, а також сам моніторинг має проводитись до семи разів на добу.

Суть функціонування неінвазивних вимірювачів глюкози полягає в аналізі за станом кровоносних судин, тобто передбачаються не прямі вимірювання. Окрім цього, існує декілька функціональних моделей, які забезпечують обчислення концентрації глюкози у крові на основі аналізу стану шкірного покриву. Для цього достатньо забезпечити контакт пристрою з ділянкою частини тіла людини.

Загалом за технологіями неінвазивного визначення рівня цукру в крові використовуються технології, які показано на рис. 1.4.

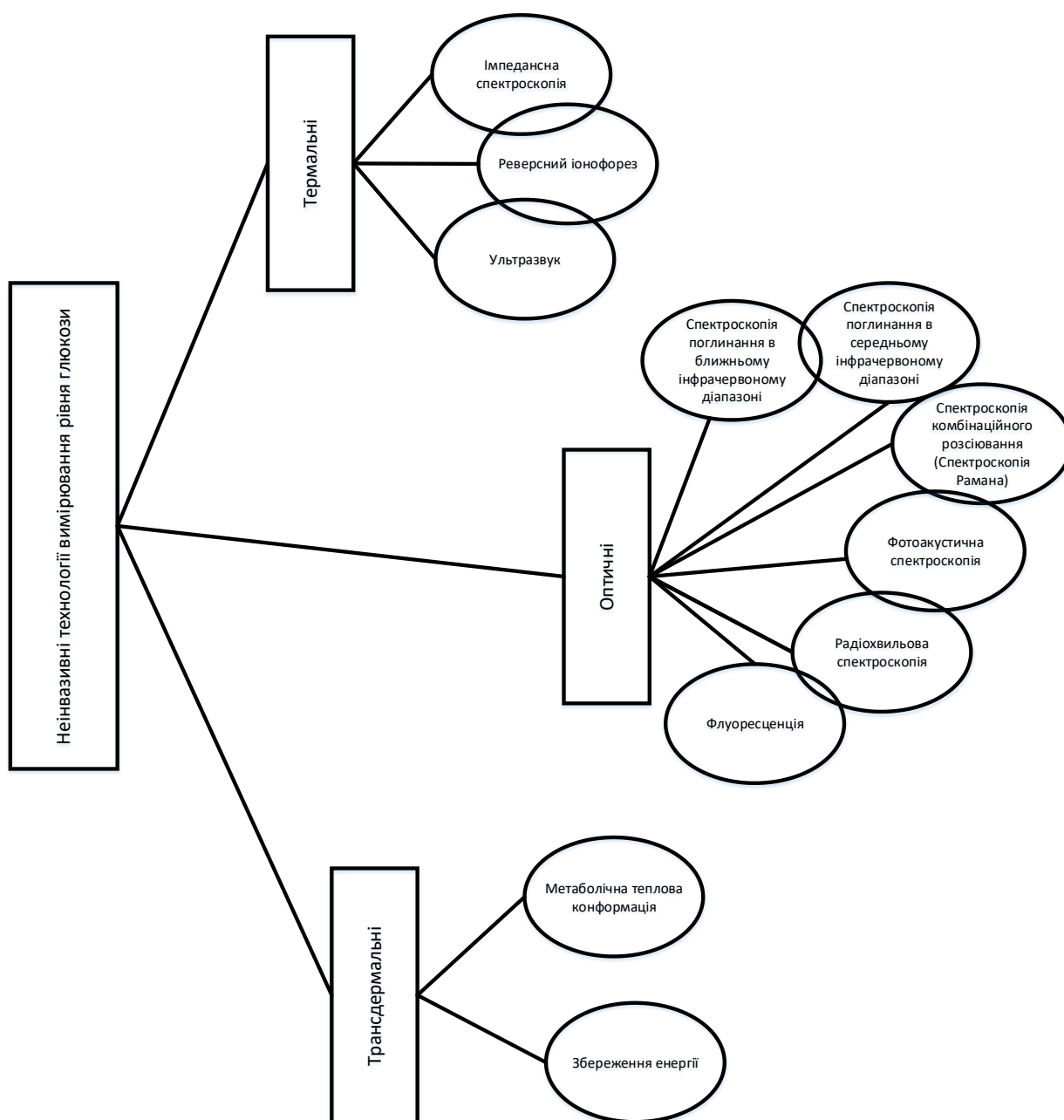


Рис. 1.4. Технологія неінвазивного визначення рівня цукру в крові

Оптичні методи використовують різні властивості світла для взаємодії з глюкозою в залежності від її концентрації. Трансдермальний метод передбачає вимірювання рівня глюкози через шкіру за допомогою електричних імпульсів або ультразвуку. Нарешті, термічні методи спрямовані на вимірювання рівня глюкози шляхом виявлення фізіологічних показників, пов'язаних з метаболічним утворенням тепла. На трансдермальні методи впливають зміни навколишнього середовища, наприклад, температура та піт [13], тоді як основним обмеженням оптичних технологій є те, що вони залежать від властивостей досліджуваної матерії, таких як колір шкіри [14].

Світло MIR поширюється лише на кілька мікрометрів і може бути застосовано для аналізу зразка крові. З іншого боку, NIR світло має більш глибоке проникнення в біологічне середовище до кількох міліметрів. NIR має потенціал для застосування для неінвазивного або мінімально інвазивного аналізу крові, навіть якщо абсорбція глюкози не така висока, як у MIR-області. Нижче наведено найпоширеніші неінвазивні методи.

При застосуванні інфрачервоної спектроскопії низької інтенсивності глюкоза дає один із найслабших сигналів поглинання в такому інфрачервоному випромінюванні на одиницю концентрації основного компонента в організмі. Вимірювання рівня глюкози за допомогою спектроскопії в ближньому інфрачервоному випромінюванні дозволяє досліджувати глибину тканин у діапазоні від 1 до 100 міліметрів із загальним зменшенням глибини проникнення зі збільшенням значення довжини хвилі. Інфрачервоне випромінювання низької інтенсивності проникає через мочку вуха, павутинку та кутикулу пальців або відбившись від шкіри.

Інша технологія для неінвазивного моніторингу рівня глюкози в крові за допомогою спектрокопа вимірює поглинання дальнього випромінювання в інфрачервоному діапазоні (FIR). Воно міститься в природному тепловому спектрі і за допомогою відповідного пристрою дозволяє вимірювати FIR поглинання, наявне у природному тепловому випромінюванні або теплі тіла. FIR-спектроскопія – єдиний тип радіаційної технології, яка не потребує зовнішнього джерела енергії.

Раманівська спектроскопія вимірює розсіяне світло, на яке впливають коливання та обертання розсіяного світла. Для розчинів крові, води, сироватки та плазми випробували різні методи Рамана. Аналітичні проблеми включають нестабільність довжин хвиль та інтенсивності лазера та помилки через інші хімічні речовини у зразку тканини та тривалий час отримання спектру [3]

Фотоакустична спектроскопія використовує оптичний промінь для швидкого нагрівання зразка та створення хвилі акустичного тиску, яку можна виміряти мікрофоном. Методи також зазнають хімічного впливу біологічних молекул, а також фізичного впливу змін температури і тиску.

Глюкоза в крові відповідає за забезпечення організму енергією. Для спектрофотометричних експериментів використовується і розвивається закон Ламберта позначення поглинання для вираження поглинання світла як функції концентрації глюкози в крові, що задається наступним рівнянням

$$OD = \log_{10} \frac{I_0}{I} = k * l * C \quad (1.1)$$

де OD – оптична щільність,

I_0 – інтенсивність падаючого світла;

I – інтенсивність проникаючого світла;

k – коефіцієнт пропорційності;

C – концентрація глюкози в крові;

l – довжина шляху світла через розчин.

Конструкція датчика

Як приклад, розглянемо систему на основі сенсора для вимірювання рівня глюкози [6]. До її складу входить ряд апаратних модулів, які формують відповідні джерела світла, кола постійної інтенсивності, схеми постійної інтенсивності, схеми детектора. Сенсор складається з емітерних світлодіодів з середньою довжиною хвилі простого світла червоного діапазону з довжиною хвилі 940 нм та 1450 нм. Порівняльне дослідження для всіх довжин хвиль виконується на спеціально

розробленому обладнанні, де кожна довжина хвилі демонструє різну поведінку при зчитуванні з відомою концентрацією глюкози під впливом різних довжин хвиль. Світлодіоди призначені для розміщення навпроти фотодіода, який виявляє світло від світлодіодів. Поглинання на кожній довжині хвилі суттєво відрізняється для різних концентрацій глюкози

Детектор загальний для всіх довжин хвиль, а його схема показана на рис. 1.5.

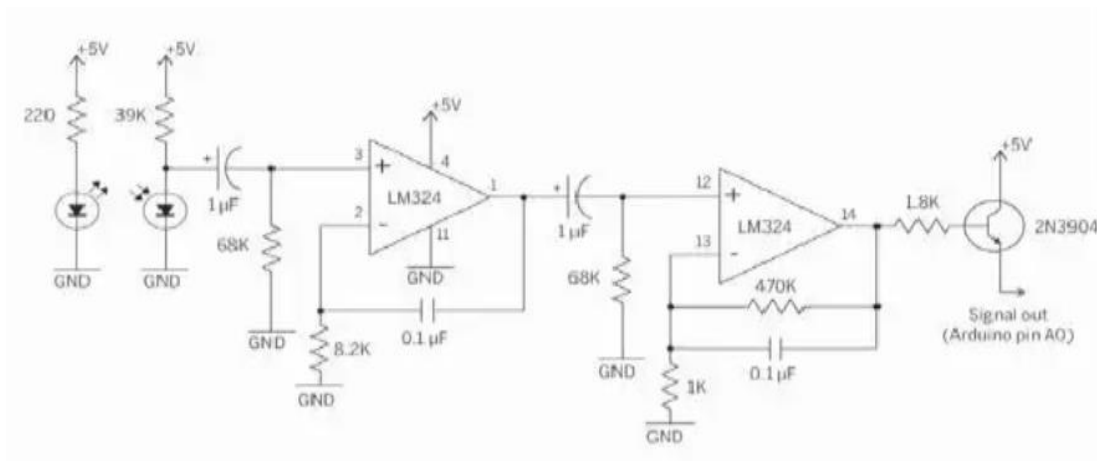


Рис. 1.5. Схема детектора довжини хвиль

Зв'язок із концентрацією глюкози, коли експерименти проводяться з різними її концентраціями, отримані з дистильованої води та цукру показано на рис. 1.6.

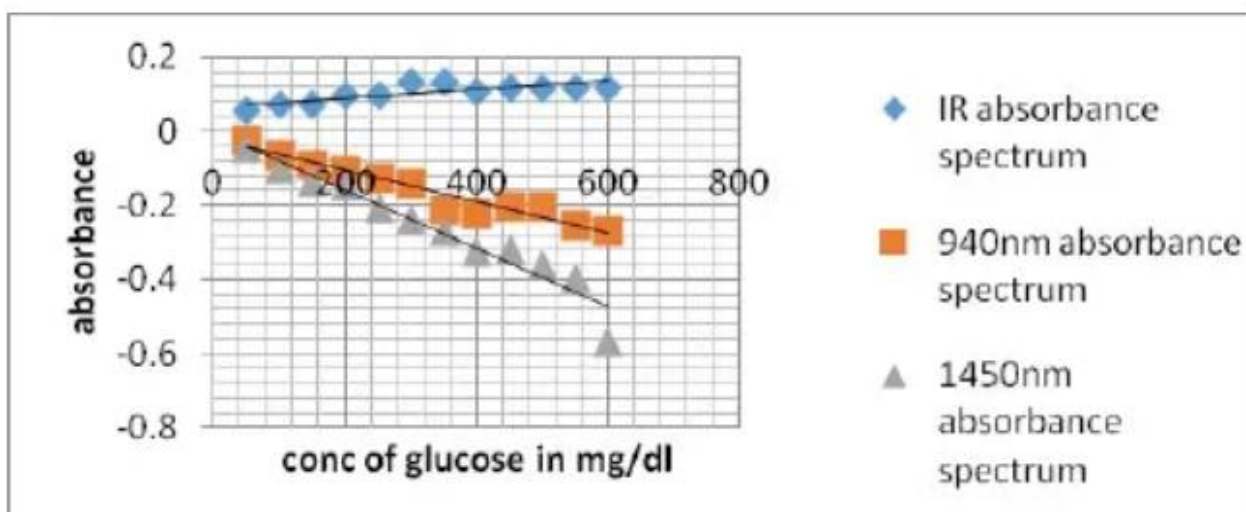


Рис. 1.6. Залежності поглинання світла при різній довжині хвилі у розчині

Цей єдиний фотодіод приймача встановлений у нижній частині затискача для пальців. Зонд прикладають до тіла пацієнта зазвичай на палець. Пропущене світло сприймається фотодіодом. Вихідна напруга фотодіода змінюється лінійно як і інтенсивність.

Оптичний датчик рівня глюкози в крові використовує довжину хвилі простого ІЧ-випромінювання 940 нм і 1450 нм. Вихідні дані отримують на цифровому накопичувальному осцилографі. Встановлюється кореляція між фактичною концентрацією рівня глюкози в крові та вимірною вихідною напругою. З результатів видно, що 940 нм і простий ІЧ дають лінійну залежність, але довжину хвилі 1450 нм можна використовувати для неінвазивного вимірювання, оскільки вона має лінійну залежність від концентрації глюкози за умови зменшення інших елементів, що викликають помилки.

1.3. Висновки до розділу

До основних результатів, отриманих у даному розділі, належать:

1. У результаті проведеного аналізу наукових досліджень і статистичних даних щодо захворюваності на цукровий діабет як в Україні, так і в світі, встановлено що дане захворювання є досить розповсюдженим і зберігає тенденції до становлення як глобальної пандемії.

2. Виявлено основні фактори впливу на розвиток і появу підвищеного рівня глюкози в крові, що в основному залежить від специфіки життєдіяльності людини, віку та пізнього звернення у лікувальні установи, що дало змогу обґрунтувати доцільність та актуальність розробки і впровадження комп'ютерних систем виявлення, аналізу та прогнозування розвитку цукрового діабету.

3. Проаналізовано існуючі підходи, методи та інструменти для визначення рівня цукру в крові людини і встановлено, що найбільш поширеними для побутового використання є пристрої інвазивного та неінвазивного вимірювання, а системи моніторингу в Україні відсутні або недосконалі, що не дає змоги

забезпечити якість обслуговування пацієнтів та виконувати прогнозування розвитку цього захворювання.

4. На основі результатів проведеного аналізу методів і способів опрацювання даних щодо рівня глюкози в крові обґрунтовано необхідність організації проекту апаратного і програмного забезпечення для розробки системи збору, аналізу і прогнозування імовірності появи і розвитку цукрового діабету.

РОЗДІЛ 2

ПОБУДОВА АРХІТЕКТУРИ ПРОГРАМНО-АПАРАТНИХ ЗАСОБІВ ОПРАЦЮВАННЯ ДАНИХ В СИСТЕМАХ МОНІТОРИНГУ РІВНЯ ЦУКРУ В КРОВІ

2.1. Структура та алгоритми роботи пристроїв вимірювання глюкози в крові

У кваліфікаційній роботі пропонується розглянути глюкометр, як вбудовану систему. В загальному випадку, комплекс програмно-апаратних засобів для опрацювання даних при визначенні і моніторингу рівня цукру в крові базуються на наступних типах систем:

1. Вбудована система – це система, де програмне забезпечення керує апаратним пристроєм і є вбудованим у цей пристрій. Проблеми у вбудованих системах, зазвичай, включають фізичний розмір, швидкість реакції, керування живленням тощо. У даному випадку, вбудована система, розглядається як програмна система для керування медичним пристроєм.

2. Інформаційна система – це система, основною метою якої є управління та надання доступу до бази даних інформації. Проблеми інформаційних систем включають безпеку, зручність використання, конфіденційність і збереження цілісності даних. Інформаційна система в даному випадку представляє собою систему медичних записів показників рівня глюкози.

3. Система збору даних на основі датчиків – це система, основна мета якої полягає в зборі даних із набору датчиків та обробці цих даних певним чином. Ключовими вимогами таких систем є надійність навіть у несприятливих умовах та доступність.

2.1.1. Система вбудованого глюкометра

Інсулінова помпа – це медична система, яка імітує роботу підшлункової залози (внутрішнього органу). Програмне забезпечення, що керує цією системою,

є вбудованою системою, яка збирає інформацію з датчика та керує насосом, який доставляє контрольовану дозу інсуліну користувачеві.

Системою користуються люди, які страждають на цукровий діабет. Діабет є відносно поширеним станом, коли підшлункова залоза людини не може виробляти достатню кількість гормону, який називається інсулін. Інсулін метаболізує глюкозу (цукор) у крові.

Традиційне лікування цукрового діабету передбачає регулярні ін'єкції генно-інженерних препаратів інсулін. Діабетики вимірюють рівень цукру в крові за допомогою зовнішнього глюкометра, а потім розраховують дозу інсуліну, яку вони повинні ввести.

Проблема цього лікування полягає в тому, що необхідний рівень інсуліну залежить не тільки від рівня глюкози в крові, а й від часу останньої ін'єкції інсуліну. Це може призвести до дуже низького рівня глюкози в крові (якщо інсуліну занадто багато) або дуже високого рівня цукру в крові (якщо інсуліну занадто мало).

Низький рівень глюкози в крові в короткостроковій перспективі є більш серйозним станом, оскільки він може призвести до тимчасового порушення роботи мозку і, зрештою, до втрати свідомості та смерті. У довгостроковій перспективі, однак, постійно високий рівень глюкози в крові може призвести до пошкодження очей, нирок і проблем з серцем.

Сучасні досягнення в розробці мініатюрних датчиків означають, що тепер можна розробляти автоматизовані системи доставки інсуліну. Ці системи контролюють рівень цукру в крові та вводять необхідну дозу інсуліну. Такі системи доставки інсуліну вже існують для лікування пацієнтів лікарні. У майбутньому, можливо, у багатьох діабетиків такі системи будуть постійно прикріплені до тіла.

Програмно керована система доставки інсуліну може працювати, використовуючи мікросенсор, вбудований в організм пацієнта, для вимірювання деяких параметрів крові, пропорційних рівню цукру. Потім це надсилається на контролер насоса. Цей контролер обчислює рівень цукру та необхідну кількість інсуліну. Потім він надсилає сигнали мініатюрному насосу, щоб доставити інсулін

через постійно прикріплену голку. На рис. 2.1 показано структуру вбудованої системи для моніторингу рівня глюкози в крові.

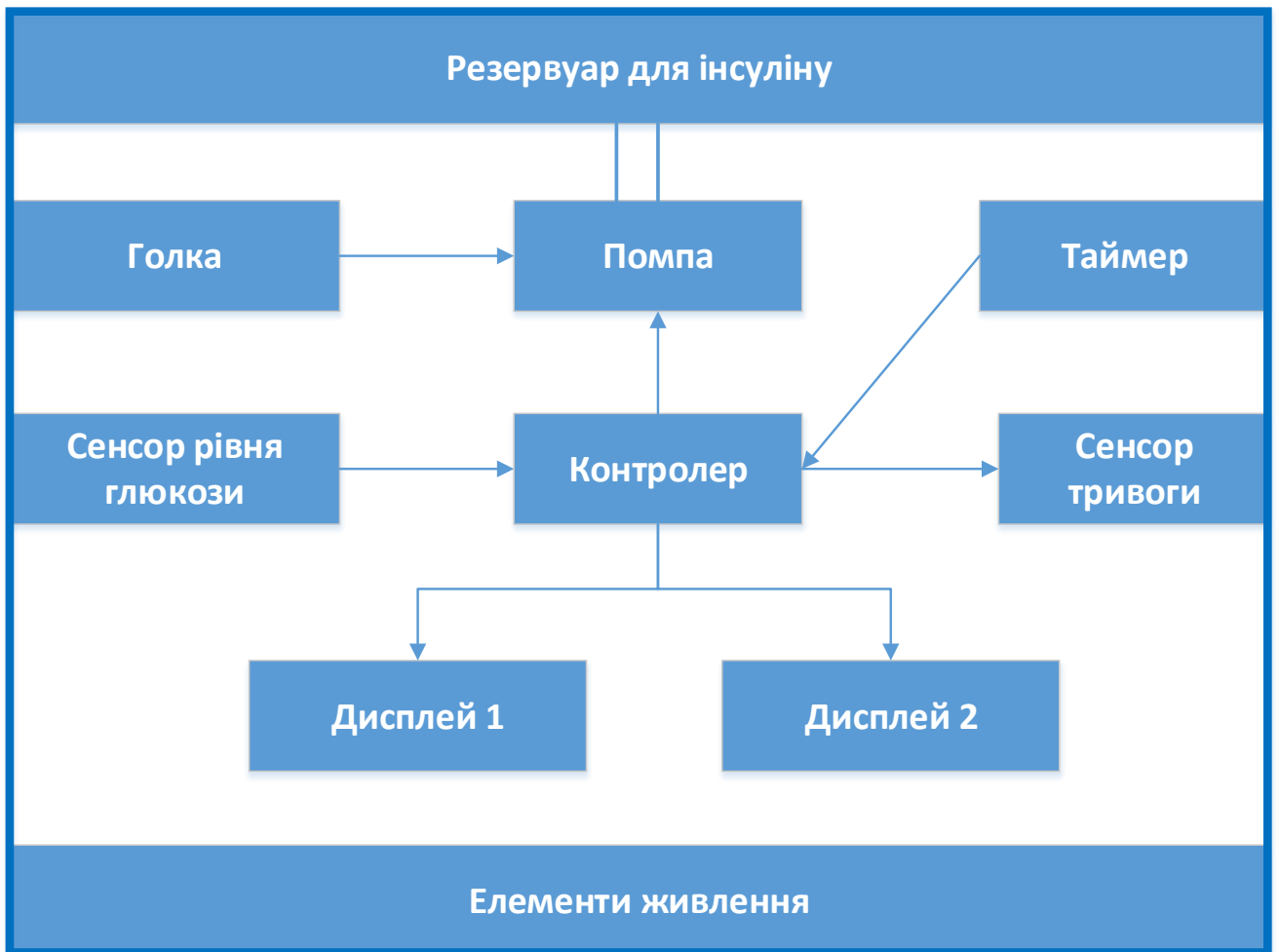


Рис. 2.1. Вбудована система контролю рівня глюкози в крові

На рис. 2.1 показані апаратні компоненти та організацію інсулінової помпи. Потрібно знати те, що датчик крові вимірює електропровідність за різних умов і що ці значення можуть бути пов'язані з рівнем цукру в крові.

Інсулінова помпа подає одну одиницю інсуліну у відповідь на один імпульс від контролера. Тому для доставки 10 одиниць інсуліну контролер посилає помпі 10 імпульсів.

На рис. 2.2 показана модель у вигляді діаграми видів діяльності UML, яка ілюструє роботу програмного забезпечення – перетворює вхідний рівень цукру в крові на послідовність команд, які керують інсуліновою помпою.

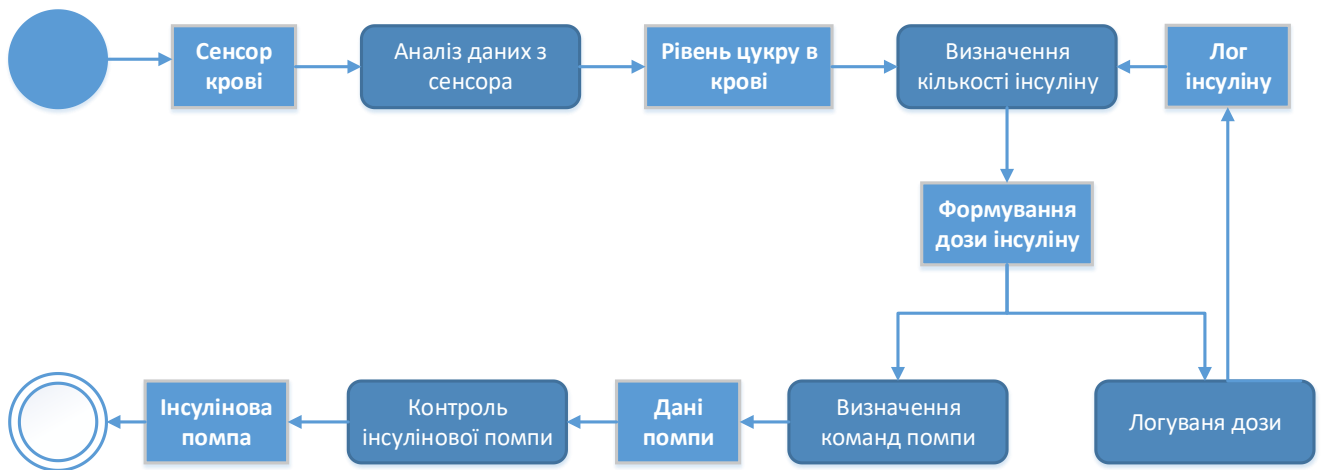


Рис. 2.2. Алгоритм функціонування системи моніторингу рівня цукру в крові

Зрозуміло, що це критична для безпеки система. Якщо насос не спрацює або не працює належним чином, це може зашкодити здоров'ю користувача або він може впасти в кому, оскільки рівень цукру в крові занадто високий або занадто низький. Таким чином, ця система має відповідати двом важливим вимогам високого рівня:

1. Система повинна бути доступна для доставки інсуліну, коли це необхідно.
2. Система повинна працювати надійно та доставляти потрібну кількість інсуліну для протидії поточному рівню цукру в крові.

Тому система повинна бути розроблена та впроваджена таким чином, щоб завжди відповідати цим вимогам.

Специфікація може бути структурована навколо об'єктів, якими керує система, функцій, які виконує система, або подій, оброблених системою. Приклад специфікації на основі форми, у цьому випадку такої, яка визначає, як розрахувати дозу інсуліну, яка буде введена, коли рівень цукру в крові знаходиться в межах безпечного діапазону, показано у табл. 2.1.

Вимоги щодо контролю дози інсуліну у системі

Стан пацієнта	Дія системи
Рівень глюкози в крові знаходиться в межах ($r_2 < r_1$), тобто є низьким	$CompDose=0$
Рівень глюкози в крові знаходиться в межах ($r_2=r_1$), тобто є нормальним	$CompDose=0$
Рівень глюкози в крові знаходиться в межах ($(r_2 < r_1) < (r_1 - r_0)$), тобто є підвищеним в межах похибки	$CompDose=0$
Рівень глюкози в крові знаходиться в межах ($(r_2 < r_1) \geq (r_1 - r_0)$), тобто є підвищеним	$CompDose=round((r_2-r_1)/4)$ If rounded result=0 then $CompDose=MinimumDose$

В якості прикладу, розглянемо програмний код, показаний на рис. 2.3, який може бути частиною реалізації системи доставки інсуліну.

```

- The insulin dose to be delivered is a function of
- blood sugar level, the previous dose delivered and
- the time of delivery of the previous dose
currentDose = computeInsulin ( ) ;
// Safety check-adjust currentDose if necessary.
// if statement 1
if (previousDose == 0)
{
    if (currentDose > maxDose/2)
        currentDose = maxDose/2 ;
}
else
    if (currentDose > (previousDose * 2) )
        currentDose = previousDose * 2 ;
// if statement 2
if ( currentDose < minimumDose )
    currentDose = 0 ;
else if ( currentDose > maxDose )
    currentDose = maxDose ;
administerInsulin (currentDose) ;

```

Рис. 2.3. Приклад можливої реалізації програмного коду для забезпечення безпеки користувача (хворого)

Код обчислює дозу інсуліну, яку потрібно ввести, а потім застосовує деякі перевірки безпеки, щоб зменшити ймовірність передозування інсуліну.

Розробка аргументу безпеки для цього коду передбачає демонстрацію того, що введена доза інсуліну ніколи не перевищує максимально безпечний рівень для однієї дози. Це встановлюється для кожного окремого користувача діабетика у відповідності з його медичними показаннями. Щоб продемонструвати безпеку, не потрібно доводити, що система доставляє «правильну» дозу, лише те, що вона ніколи не доставляє передозування пацієнту.

Алгоритм працює, виходячи з припущення, що `maxDose` є безпечним рівнем для цього користувача системи. Щоб побудувати аргумент безпеки, потрібно визначити предикат, який визначає небезпечний стан, яким є поточна доза `maxDose`.

Після цього потрібно продемонструвати, що всі програмні шляхи призводять до суперечності цього небезпечного твердження. Якщо це так, небезпечна умова не може бути істинною.

Якщо це зробити, то можна з впевненістю сказати, що програма не обчислить небезпечну дозу інсуліну. Для цього варто структурувати та представити аргументи безпеки графічно, як показано на рис. 2.4.

Щоб побудувати структурований аргумент для того, щоб програма не створювала небезпечних обчислень, спочатку визначаються усі можливі шляхи у коді, які можуть призвести до потенційно небезпечного стану. Тут потрібно працювати у зворотному напрямку від цього небезпечного стану та розглядати останнє присвоєння усім змінним стану на кожному шляху, що веде до цього небезпечного стану.

Якщо існує можливість показати, що жодне зі значень цих змінних не є небезпечним, то початкове припущення (те, що обчислення є небезпечним) є неправильним.

Робота у зворотному напрямку важлива, оскільки це означає, що можна ігнорувати всі стани, окрім кінцевих станів, які призводять до умови виходу з коду. Попередні значення не мають значення для безпеки системи. У цьому прикладі все,

що потрібно мати на увазі, це набір можливих значень `currentDose` безпосередньо перед виконанням методу `administerInsulin`.

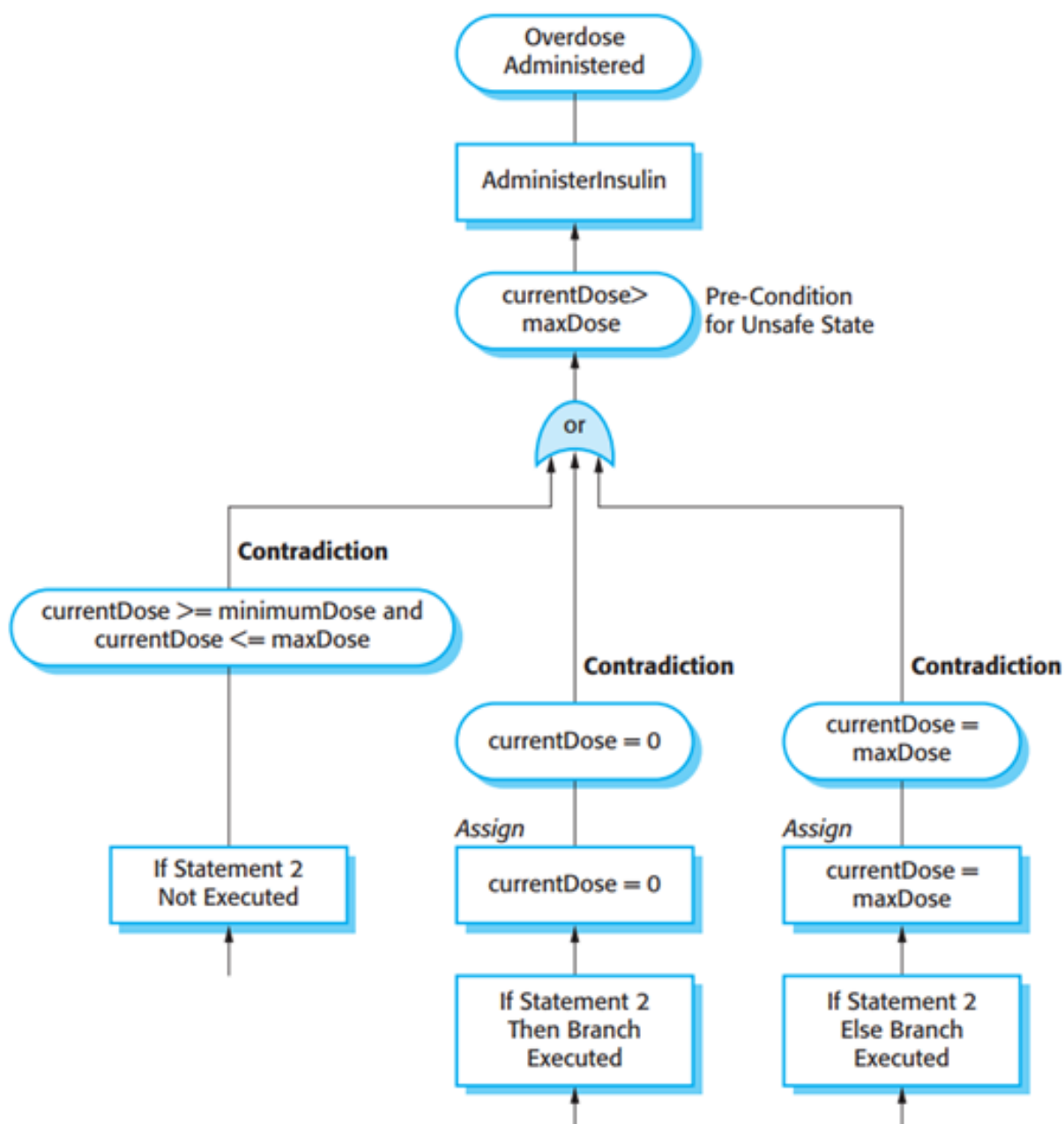


Рис. 2.4. Приклад контролю безпеки при введенні дози інсуліну

Як видно з рис. 2.4, можна ігнорувати обчислення (оператор `if 1`) в аргументі безпеки, оскільки їхні результати перезаписуються в наступних операторах програми. В аргументі безпеки, показаному на малюнку 2.5, є три можливі програмні шляхи, які призводять до виклику методу `administerInsulin`.

У даному випадку потрібно показати, що кількість доставленого інсуліну ніколи не перевищує максимальну дозу. Для цього розглядаються всі можливі програмні шляхи введення інсуліну:

1. Жодна гілка оператора `if 2` не виконується. Це може статися, лише якщо поточна доза більша або дорівнює мінімальній дозі та менша або дорівнює максимальній дозі. Це постумова — твердження, яке є істинним після виконання оператора.

2. Виконується гілка `then` оператора `if 2`. У цьому випадку виконується встановлення поточної дози на нуль. Тому його пост-умова поточна доза 0.

3. Виконується гілка `else-if` оператора `if 2`. У цьому випадку виконується налаштування призначення `currentDose` на `maxDose`. Отже, після виконання цього оператора відомо, що пост-умова це `currentDose maxDose`.

У всіх трьох випадках постумови суперечать небезпечній попередній умові, згідно з якою введена доза перевищує `maxDose`. Тому можна стверджувати, що обчислення є безпечним.

2.1.2. Система збору та опрацювання показників рівня глюкози в крові

Інформаційна система пацієнта для опрацювання показників його здоров'я — це медична інформаційна система, яка зберігає інформацію про пацієнтів, які мають проблеми зі здоров'ям, і одержують певний вид лікування.

Більшість пацієнтів з діагностованим діабетом не потребують спеціального лікування в стаціонарі, але їм необхідно регулярно відвідувати спеціалізовані клініки, де вони можуть зустрітися з лікарем, який добре знає їхні проблеми.

Щоб пацієнтам було легше відвідувати та контролювати стан здоров'я необхідно забезпечити віддалений спосіб комунікації та зберігання даних про рівень цукру в крові. Для цього, у найбільш розвинутих країнах створюють інформаційні системи, призначені для використання в клініках. Зазвичай, вони використовують централізовану базу даних з інформацією про пацієнтів, але також є варіанти реалізації під веб, що забезпечує зручність і доступність з будь-якого ПК

чи мобільного пристрою. Тому доступ до ресурсу можна отримати і без наявного безпечного підключення до мережі.

Коли локальні системи мають безпечний доступ до мережі, вони використовують інформацію про пацієнта в базі даних, але можуть завантажувати та використовувати локальні копії записів пацієнтів, коли вони відключені.

Система не є повною системою медичних записів, тому не зберігає інформацію про інші захворювання. Однак він може взаємодіяти та обмінюватися даними з іншими клінічними інформаційними системами.

На рис. 2.5 проілюстровано загальну концепцію організації систем для збору та аналізу медичної інформації.

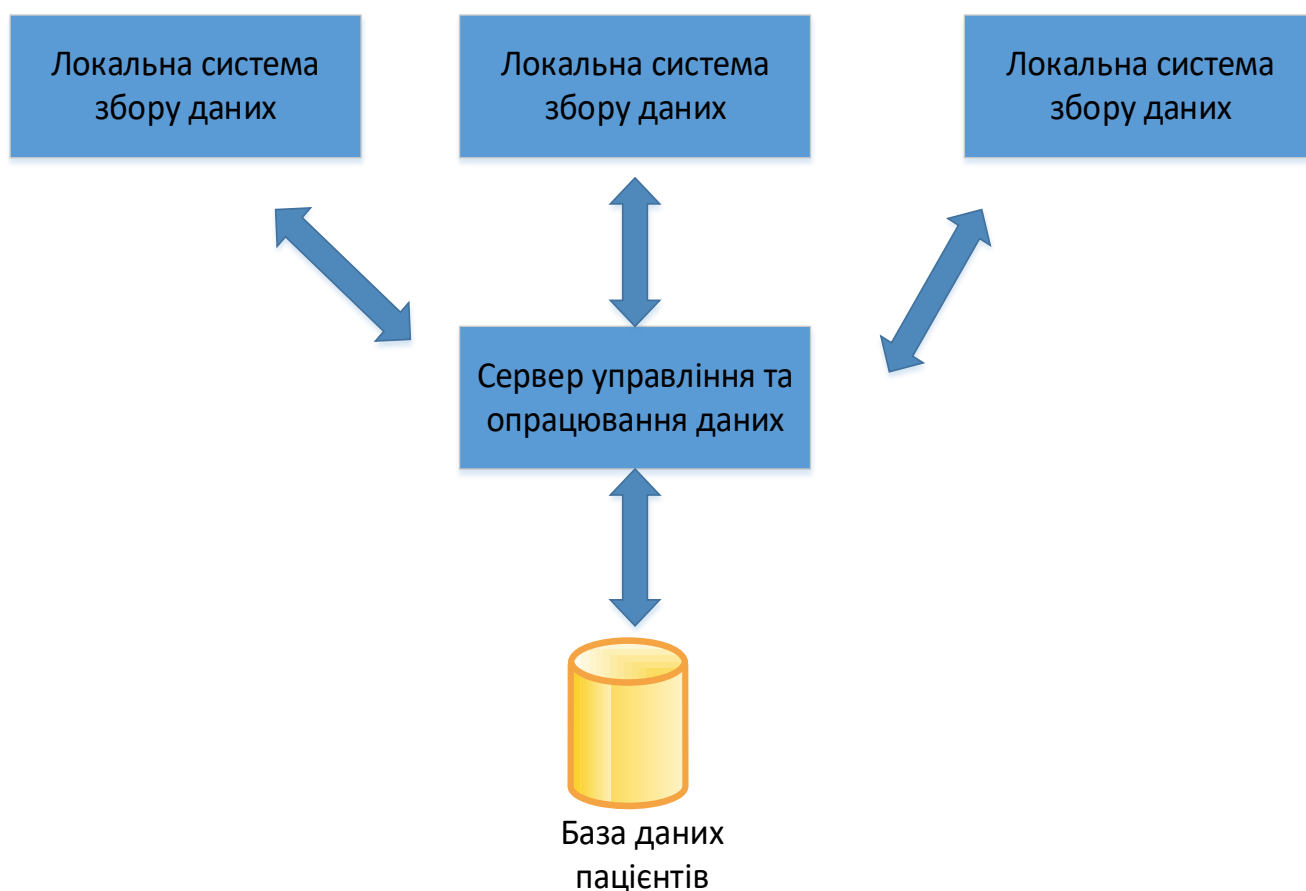


Рис. 2.5. Концептуальна схема організації системи збору та опрацювання медичних даних

На рис 2.5 проілюстровано організацію системи, яка має дві загальні цілі:

1. Згенерувати управлінську інформацію, яка дозволить керівникам медичних служб оцінити результати діяльності щодо місцевих цілей.

2. Своєчасно надавати медперсоналу інформацію для підтримки лікування пацієнтів.

Природа проблем, пов'язаних зі здоров'ям людей, хворих цукровим діабетом така, що пацієнти часто неорганізовані, тому можуть пропускати прийоми, навмисно або випадково втрачати рецепти та ліки, забути інструкції та висувати необґрунтовані вимоги до медичного персоналу.

У меншій частині випадків вони можуть становити небезпеку для них самих або для інших людей. Пацієнти можуть регулярно змінювати адресу на довгостроковій або короткостроковій основі. Там, де пацієнти є «небезпечними», їх, можливо, потрібно «розділити» - помістити в безпечну лікарню для лікування та спостереження.

Користувачами системи є клінічний персонал, такий як лікарі, медсестри та відвідувачі (медсестри, які відвідують людей вдома, щоб перевірити їх лікування).

Немедичні користувачі включають реєстраторів, які призначають зустрічі, працівників медичної документації, які ведуть систему обліку, та адміністративного персоналу, який створює звіти.

Система використовується для запису інформації про пацієнтів (ім'я, адреса, вік, найближчі родичі тощо), консультації (дата, огляд лікаря, суб'єктивні враження пацієнта тощо), стани та лікування.

Звіти формуються через регулярні проміжки часу для медичного персоналу та керівників органів охорони здоров'я. Як правило, звіти для медичного персоналу зосереджуються на інформації про окремих пацієнтів, тоді як звіти керівництва є анонімними та стосуються умов, витрат на лікування тощо. Основними характеристиками системи є:

1. Управління індивідуальним обслуговуванням.

Клініцисти можуть створювати записи для пацієнтів, редагувати інформацію в системі, переглядати історію пацієнтів тощо. Система підтримує зведення даних,

щоб лікарі, які раніше не зустрічалися з пацієнтом, могли швидко дізнатися про ключові проблеми та методи лікування, які були призначені.

2. Моніторинг пацієнтів.

Система регулярно відстежує записи пацієнтів, які залучені до лікування, і видає попередження у разі виявлення можливих проблем. Тому, якщо пацієнт деякий час не звертався до лікаря, може бути винесено попередження. Одним з найважливіших елементів системи моніторингу є відстеження пацієнтів, які були розділені, і забезпечення того, щоб законодавчо необхідні перевірки проводилися в потрібний час.

3. Адміністративна звітність.

Система створює щомісячні звіти, що показують кількість пацієнтів, які пройшли лікування в кожній клініці, кількість пацієнтів, які увійшли та вийшли з системи допомоги, кількість пацієнтів, які перебувають у відділенні, призначені ліки та їх вартість тощо.

Два різних закони впливають на медичну систему. Це закони про захист даних, які регулюють конфіденційність особистої інформації та закони про здоров'я, які регулюють таємність хвороб пацієнтів.

Однією з цілей системи моніторингу та опрацювання медичних даних є забезпечення того, щоб співробітники завжди діяли відповідно до закону, а їхні рішення реєструвалися.

Як і в усіх медичних системах, конфіденційність є важливою системною вимогою. Важливо, щоб інформація про пацієнтів була конфіденційною і ніколи не розкривалася нікому, крім уповноваженого медичного персоналу та самого пацієнта.

Медичні інформаційні системи також є критичним з точки зору безпеки. Деякі, наприклад, психічні захворювання викликають у пацієнтів суїцидальні настрої або становлять небезпеку для інших людей. По можливості система повинна попереджати медичний персонал про потенційно суїцидальних або небезпечних пацієнтів.

Загальний дизайн системи повинен враховувати вимоги конфіденційності та безпеки. Система повинна бути доступною, коли це необхідно, інакше безпека може бути порушена, і може бути неможливо призначити правильні ліки пацієнтам. Тут є потенційний конфлікт — конфіденційність найлегше підтримувати, коли є лише одна копія системних даних. Однак, щоб забезпечити доступність у разі збою сервера або при відключенні від мережі, слід зберігати декілька копій даних, що зумовлює необхідність реалізації розподіленої системи.

2.2. Можливі рішення щодо забезпечення розподіленості системи збору та опрацювання медичних даних

Для забезпечення оптимізації побудови системи збору та опрацювання даних про рівень глюкози в крові необхідно дослідити особливості проектування та функціонування розподілених комп'ютерних систем з можливістю зберігання та маніпулювання базами даних. Важливими факторами при цьому є визначення подібності та однотипності систем управління базами даних, способів і засобів логічного їх розподілу за вузлами утвореної мережі та формування запитів/відповідей з врахуванням фрагментів БД.

Загальноприйнятим є підхід і забезпечення властивості фрагментів бази даних таким чином, щоб дані розміщались якнайближче до тих вузлів, де вони найбільше потрібні. При цьому, частина або фрагмент структурованих даних при використанні реляційного підходу представляється у вигляді окремої таблиці або сукупності таких таблиць.

У кваліфікаційній роботі при організації системи опрацювання даних щодо моніторингу рівня глюкози в крові пропонується застосувати підхід, коли локальний фрагмент у вигляді таблиці створюється на локальних вузлах шляхом одержання даних з глюкометра. Після цього формується запит до центрального вузла і при потребі воно передається у глобальне сховище. Ще один спосіб організації розподіленої системи передбачає застосування єдиної схеми БД для усіх

вузлів мережі, включно з центральним, а логіка функціонування вузла керування забезпечує підтримку актуальності даних у часі.

Фрагментом даних, для першого варіанту побудови системи, є результат, який одержаний як відповідь на запит до локального хоста. У такому випадку потрібно вирішити проблему оптимальності розподілу реляційних відношень за вузлами розподіленої системи.

Суть та найбільш важлива задача фрагментації даних полягає у представленні деякої таблиці як множини частинок у такому вигляді, який відповідає параметрам ефективності. Такими параметрами можуть виступати час з'єднання з вузлом, час відгуку надсилання відповіді, обсяг оперативної пам'яті, кількість одночасних звернень та ряд інших.

Загальними правилами при створенні фрагментів таблиці є відповідність критеріям коректності, що обумовлює повноту схеми бази даних, відсутність перетинів та здатність бути реконструйованими.

Повноту фрагментації прийнято трактувати як здатність забезпечити і встановити відношення кожного запису даних до одного з фрагментів таблиць бази даних.

Якщо кожен запис відношення бази даних належить одному і тільки одному фрагменту, то така фрагментація відповідає критерію відсутності перетинів.

У випадку наявності деякої операції або функціоналу, які можуть відтворити дані з утворених фрагментів до вихідної таблиці, то такий спосіб фрагментації відповідає критерію здатності до реконструювання.

У теорії і на практиці виділяють горизонтальний, вертикальний та змішаний типи фрагментації, які дають змогу забезпечити ефективність розподілу даних за вузлами мережі при використанні реляційного підходу.

Для того, щоб виконати горизонтальну фрагментацію потрібно забезпечити такий розподіл даних за рядками таблиці із застосуванням визначеної умови (предикат з оператором або кількома операторами селекції), щоб структура таблиці не зазнала змін, в відмінними були тільки екземпляри фрагментів. На рис. 2.6 продемонстровано приклад реалізації горизонтальної фрагментації.

ID_Product	ID_Category	ID_Measur...	Title	Price
1	1	2	Milk 3,4%	25
2	1	2	Milk 2,5%	22
3	1	2	Milk 1,5%	20

1<=ID_Product<3

ID_Prod...	ID_Categ...	ID_Measurem...	Title	Price
4	1	3	Sour cream 15%	26
5	1	3	Sour cream 20%	30
6	1	3	Sour cream 30%	35
7	1	2	Yougurt 1	18
8	1	2	Yougurt 2	18

ID_Product>3

Рис. 2.6. Горизонтальна фрагментація

За істинності умови щодо відновлення початкового стану таблиці шляхом застосування операції об'єднання множин до розподілених фрагментів, вважається що такий розподіл є повним [9]

$$R = \bigcup_{i=1,k} \sigma_{F_i}(R) \quad (2.2)$$

R – таблиця, яку фрагментували;

σ_{F_i} – сукупність умов фрагментації таблиці.

Різновидом горизонтальної фрагментації є породжена горизонтальна фрагментація. Цей різновид утворюється за наявності у схемі бази даних типу зв'язку «один-до-багатьох», де бере участь вихідне фрагментоване відношення і його поділ спричиняє породжену горизонтальну фрагментацію у кінці зв'язку, тобто у таблиці із зовнішнім ключем.

Як правило, за наявності такого різновиду горизонтальної фрагментації, таблиці пов'язані зв'язком «один-до-багатьох» рекомендується розміщати на

одному і тому ж вузлі розподіленої мережі даних. Це дозволить забезпечити ефективність і продуктивність опрацювання запитів до БД.

Наступний тип декомпозиції реляційних відношень передбачає формування фрагментів шляхом застосування операції проєкції до таблиці. Це є вертикальна фрагментація.

Необхідною умовою забезпечення здатності до реконструкції вихідної таблиці при використанні вертикальної фрагментації є наявність у всіх фрагментах первинного ключа. Застосувавши операцію з'єднання за рівністю ключів проєкцій можна відтворити вихідну таблицю.

Також для вертикальної декомпозиції можна застосовувати дещо інший метод фрагментації, який полягає у використанні штучних індексів або по-іншому ідентифікаторів записів таблиці.

В якості прикладів, на рис. 2.7 і рис. 2.8 показано відповідно вихідну таблицю та її фрагменти при застосуванні вертикальної фрагментації.

ID_Product	ID_Category	ID_Measurement	Title	Price
1	1	2	Milk 3,4%	25
2	1	2	Milk 2,5%	22
3	1	2	Milk 1,5%	20
4	1	3	Sour cream 15%	26
5	1	3	Sour cream 20%	30
6	1	3	Sour cream 30%	35
7	1	2	Yougurt 1	18
8	1	2	Yougurt 2	18

Рис. 2.7. Приклад вихідної таблиці для застосування вертикальної фрагментації

Фрагмент 1			Фрагмент 2		
ID_Product	ID_Category	ID_Measurement	ID_Product	Title	Price
1	1	2	1	Milk 3,4%	25
2	1	2	2	Milk 2,5%	22
3	1	2	3	Milk 1,5%	20
4	1	3	4	Sour cream 15%	26
5	1	3	5	Sour cream 20%	30
6	1	3	6	Sour cream 30%	35
7	1	2	7	Yougurt 1	18
8	1	2	8	Yougurt 2	18

Рис. 2.8. Фрагменти вихідної таблиці після декомпозиції

Коректність виконання вертикальної декомпозиції математично може бути описана за допомогою наступного виразу [2]

$$R = \bigcup_{i=1,n} A_{R_i} \quad (2.3)$$

R – вихідна таблиця для фрагментації;

A_{R_i} – сукупність фрагментів, які одержано після декомпозиції з використанням операції проєкції.

Вираз (2.3) відображає властивість повноти фрагментації відношень, одержаних у результаті операції проєкції над стовпцями вихідної таблиці. Якщо для утворених фрагментів застосувати операцію з'єднання за рівністю значень первинних і зовнішніх ключів, то виконається умова здатності до реконструкції.

До переваг вертикальної декомпозиції належить:

- здатність забезпечити таку фрагментацію таблиці, що її частини будуть розміщені на тих вузлах, на яких вони найчастіше використовуються;
- підтримка паралельного опрацювання фрагментів відношення;
- здатність забезпечити встановлений рівень ефективності з'єднання за значеннями ідентифікаторами кортежів.

У випадку, коли до таблиці застосувати одночасно і горизонтальну, і вертикальну фрагментації одержують тип змішаної декомпозиції.

Графічно, представлення змішаної фрагментації проілюстровано на рис. 2.9, де показано послідовність застосування відповідних операцій вертикальної і горизонтальної декомпозиції.

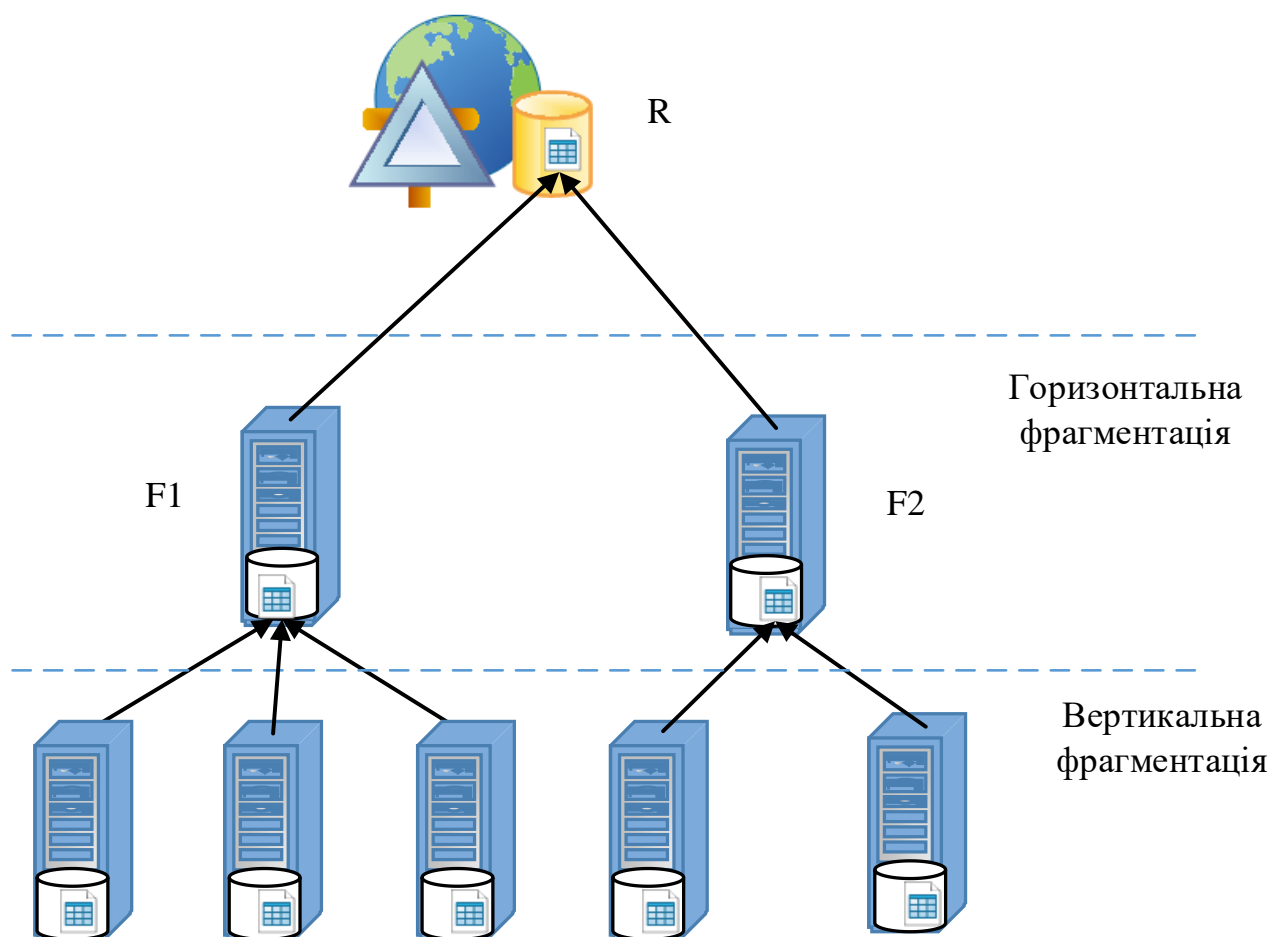


Рис. 2.9. Змішана декомпозиція

Критеріями якості виконання змішаної фрагментації виступають:

- продуктивність передачі інформації;
- час реакції на запити і формування відповіді;
- ціна щодо використання дискового простору при зберіганні даних як на окремо взятому локальному вузлі, тау і на вузлі управління;
- ліміти щодо використання ресурсів (об'єм пам'яті, характеристики інфраструктури, процесорний час і т.д.).

Вхідними даними при проектуванні розподіленої комп'ютерної системи щодо опрацювання даних при захворюваності на цукровий діабет належать:

- схема і розмір локальних фрагментів на вузлах розподіленої системи;
- потужність інструментів збору, аналізу та маніпулювання даними;
- частота звернень щодо агрегації інформації у вузлі управління системою;
- характеристики каналів обміну даними між локальними і центральним хостом системи.

При проектуванні розподіленої медичної системи збору та опрацювання даних про рівень глюкози в крові запропоновано використовувати вертикальну декомпозицію з підтримкою можливості горизонтальної фрагментації таблиць у локальних хостах системи. Проте для забезпечення оптимальності реалізації такої системи потрібно обґрунтувати способи і механізми управління одночасністю і паралельністю доступу до даних, інструменти збору та опрацювання інформації на головному вузлі управління, математично описати структуру системи в цілому.

2.3. Побудова моделі розподіленої архітектури системи збору та опрацювання даних у системах моніторингу рівня цукру в крові

Основними факторами, які потрібно врахувати при побудові моделі та проектуванні архітектури системи є кількість хостів на яких буде розміщено систему, вид фрагментації відношень для зберігання даних і їх розташування, різновид ПЗ, особливості каналів передачі даних, способи та інструменти управління інформацією із застосуванням підходу ведення транзакційних журналів. Передбачається, що транзакції можуть виконуватись як на локальних хостах системи так і на центральному вузлі управління (рис. 2.10).

Враховуючи те, що база даних інформації про пацієнтів з відповідними даними про рівень глюкози в крові та певною часовою міткою, спочатку потрапляє у локальні вузли (фрагменти), то доцільним є створення механізмів агрегації для центрального вузла управління системою. Такий механізм забезпечує трансформацію самих даних або журналу транзакцій на відповідну структуру бази

даних. Дані, які агрегуються представляють собою первинні і зовнішні ключі, особисті дані пацієнта, рівень глюкози у крові і час проведення вимірювання.

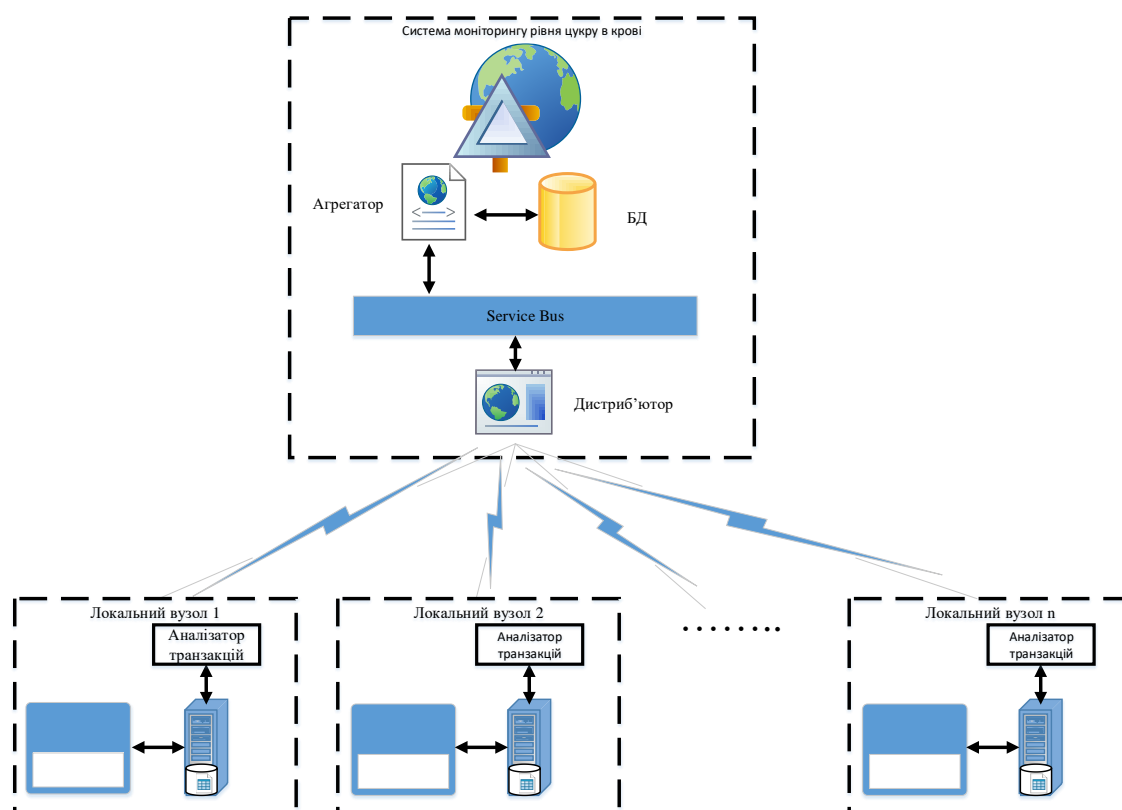


Рис. 2.10. Архітектура системи збору та опрацювання даних моніторингу рівня цукру в крові

В якості компоненту, що забезпечує збір і передачу даних використовується сервісна шина, при цьому існує можливість її застосування як на рівні локального, так і глобального хостів. Якщо відбувається зміна інформації у базі даних центрального хоста, то записи чи транзакції за допомогою агрегатора передаються до шини. Вона може виконувати роль балансувальника навантаження і забезпечувати паралельність опрацювання транзакцій. Далі управління від сервісної шини передається дистриб'ютору, функції якого полягають у забезпеченні принципу примусової часткової реплікації та надсилання транзакцій у визначені локальні хости розподіленої системи опрацювання даних щодо рівня глюкози в крові.

Локальні вузли містять спеціальні аналізатори збору транзакцій для оновлення інформації у вузлах на яких вони знаходяться. Окрім цього, аналізатори на запит дистриб'ютора можуть передавати журнал транзакцій того вузла, за яким вони закріплені у системі.

Математично, архітектуру системи моніторингу рівня цукру в крові можна, представити у вигляді елементів теорії множин, як кортеж, компонентами якого є вузли мережі різного призначення, як показано нижче

$$ADS = \langle GNode, LNode, Relation(GNode, LNode) \rangle \quad (2.4)$$

GNode – центральний вузол управління, що містить не фрагментовану базу даних;

LNode – локальні хости з фрагментованими відношеннями;

Relation(GNode, LNode) – зв'язок між центральним і локальними вузлами, що описуються відповідною парою ідентифікаторів.

LNode можна розглядати як сукупність локальних вузлів:

$$LNode = \{ lNode_i \} \quad (2.5)$$

lNode_i – хости з розміщеними на них частинами бази даних, $i = 1..N$, N - кількість хостів за якими виконано декомпозицію бази даних або фрагменти таблиць.

Центральний вузол управління представляється за допомогою множини структурних компонентів у вигляді

$$GNode = \{ Aggr, ServBus, Distr \} \quad (2.6)$$

Aggr – механізм опрацювання транзакцій;

ServBus – шина у вигляді сервісної шини;

Distr – дистриб'ютор.

Структурно локальний хост представляється у вигляді наступного запису

$$lNode_i = \{ Analyzer_i, DB_i, Transcat_{ij} \} \quad (2.7)$$

Analyzer_i – аналізатор стану БД на і-му локальному хості;

DB_i – елемент або база даних на і-му хості;

Transcat_{ij} – множина (журнал) транзакцій бази даних і-го хоста.

Сукупність операцій у вигляді транзакції для кожного локального хоста представляється у вигляді елементів множини: операція транзакції, час початку і час закінчення її виконання

$$Transcat_{ij} = \{ TBody_{ij}, StartTime_{ij}, StopTime_{ij}, TState_{ij} \} \quad (2.8)$$

TBody_{ij} – вміст j-ої транзакції в і-му хості;

StartTime_{ij} – часова мітка початку транзакції;

StopTime_{ij} – часова мітка завершення транзакції;

TState_{ij} – мітка успішності транзакції;

Запит від центрального вузла управління системи до будь-якого локального вузла, а також його вміст можна зобразити у наступному вигляді

$$Req = \{ GNode, lNode_j, Req_{ij} \} \quad (2.9)$$

Req_{ij} – запит від центрального вузла управління *GNode* до хоста *lNode_j*.

Математичне представлення процесу передачі даних від локального до центрального вузла керування системою зображено аналогічно до виразу (2.9)

$$Resp = \{ lNode_j, Resp_{ji}, GNode \} \quad (2.7)$$

$Resp_{ji}$ – відповідь хоста $lNode_j$ до центрального вузла управління $GNode$.

У результаті формального опису компонентів системи опрацювання даних при моніторингу рівня цукру спроектовано її архітектуру, що відображає механізми і запити/відповіді при комунікації центрального вузла управління з локальними хостами.

2.4. Висновки до розділу

Основні наукові та практичні результати, які отримані у даному розділі полягають в наступному:

1. Визначено і запропоновано основні компоненти апаратного і програмного забезпечення для опрацювання даних у системах моніторингу рівня цукру в крові, які містять як елементи вбудованих систем, так і комплексних інформаційних систем управління, що дало змогу забезпечити визначений ступінь функціональності та ефективності при прогнозуванні і розвитку цукрового діабету.

2. Розроблено алгоритми функціонування глюкометра та глобальної інформаційної системи управління медичними даними, які в комплексі становлять систему цілодобового моніторингу та керування рівнем глюкози в крові пацієнта, що дає змогу забезпечувати збір, опрацювання та прогнозування появи чи розвитку цукрового діабету.

3. Обґрунтовано можливі рішення щодо забезпечення розподіленості системи моніторингу та опрацювання даних захворюваністю цукровим діабетом, що дало змогу встановити оптимальний шлях побудови таких систем із застосуванням підходу змішаної фрагментації на вузлах мережі розподілу та управління навантаженням на основі програмного балансувальника.

4. Побудовано та математично представлено концептуальну модель розподіленої архітектури системи збору та опрацювання даних для моніторингу рівня цукру в крові, що включає в себе сукупність локальних і центрального вузла управління і дає змогу обмінюватися повідомленнями та виконувати прогнозування розвитку хвороби. Основна функція локальних вузлів у розподіленій системі полягає у зборі даних з глюкометрів та передачі їх у центральний вузол управління. Найбільш важлива функція центрального вузла управління полягає в агрегації даних від локальних вузлів за допомогою дистриб'ютора транзакцій, сервісної шини та агрегатора, що дозволяє забезпечувати цілісність даних та подальше прогнозування рівня цукру в крові.

РОЗДІЛ 3

ПРОГРАМНА МОДЕЛЬ ПРОГНОЗУВАННЯ РОЗВИТКУ ЦУКРОВОГО ДІАБЕТУ

3.1. Аналіз даних для прогнозування розвитку цукрового діабету

В контексті аналізу рівня глюкози в крові розрізняють підтипи діабету або схильності до нього. Ще раз звернемося до означення «діабет — це хронічний (тривалий) стан здоров'я, який впливає на те, як організм людини перетворює їжу в енергію». Організм розщеплює більшу частину їжі, на цукор (глюкозу) і викидає його в кров. Коли рівень цукру в крові підвищується, це сигналізує підшлунковій залозі виділяти інсулін. Інсулін діє як ключ, щоб пропускати цукор крові в клітини організму для використання в якості енергії.

При діабеті організм не виробляє достатньо інсуліну або не може використовувати його належним чином. Коли інсуліну недостатньо або клітини перестають реагувати на інсулін, занадто багато глюкози залишається в крові. З часом це може спричинити серйозні проблеми зі здоров'ям, такі як хвороби серця, втрата зору та захворювання нирок.

Існує 3 типи діабету, і у відкритих даних будемо мати справу з діабетом 2 типу. Цукровий діабет має 2 стадії: «Переддіабет» та «Цукровий діабет».

У роботі буде побудована модель для прогнозування діабету у пацієнтів. Для початку проаналізуємо дані, на основі яких будуть проводитись дослідження і побудова моделі та алгоритму прогнозування захворюваності на діабет 2-го типу. В якості вхідних даних використано дата сет «diabetes _ 012 _ health _ indicators _ BRFSS2015.csv», що містить 253 680 відповідей. Цільова змінна містить 3 класи:

- 0 – для людей, у яких відсутній підвищений рівень глюкози у крові;
- 1 – для людей, які мають схильність до діабету (переддіабетичний стан);
- 2 – для людей з діабетом.

У даних відображено 21 характеристику, що впливає на розвиток діабету а самі дані є незбалансованими, тобто кількість осіб, які відносяться до різних класів є різною. Фрагмент датасету показано на рис. 4.4.

# Diabetes...	# HighBP	# HighChol	# CholCheck	# BMI	# Smoker	# Stroke
0.0	1.0	1.0	1.0	40.0	1.0	0.0
0.0	0.0	0.0	0.0	25.0	1.0	0.0
0.0	1.0	1.0	1.0	28.0	0.0	0.0
0.0	1.0	0.0	1.0	27.0	0.0	0.0
0.0	1.0	1.0	1.0	24.0	0.0	0.0
0.0	1.0	1.0	1.0	25.0	1.0	0.0
0.0	1.0	0.0	1.0	30.0	1.0	0.0
0.0	1.0	1.0	1.0	25.0	1.0	0.0
2.0	1.0	1.0	1.0	30.0	1.0	0.0
0.0	0.0	0.0	1.0	24.0	0.0	0.0
2.0	0.0	0.0	1.0	25.0	1.0	0.0
0.0	1.0	1.0	1.0	34.0	1.0	0.0
0.0	0.0	0.0	1.0	26.0	1.0	0.0
2.0	1.0	1.0	1.0	28.0	0.0	0.0
0.0	0.0	1.0	1.0	33.0	1.0	1.0
0.0	1.0	0.0	1.0	33.0	0.0	0.0
0.0	1.0	1.0	1.0	21.0	0.0	0.0
2.0	0.0	0.0	1.0	23.0	1.0	0.0
0.0	0.0	0.0	0.0	23.0	0.0	0.0
0.0	0.0	1.0	1.0	28.0	0.0	0.0
0.0	1.0	1.0	1.0	22.0	0.0	1.0
0.0	1.0	1.0	1.0	38.0	1.0	0.0

Рис. 3.1. Частина даних для прогнозування розвитку цукрового діабету

Для побудови моделі прогнозування розвитку цукрового запропоновано скористатися мовою програмування Python та набором відкритих бібліотек машинного навчання. На рис. 3.2 наведено імпорт необхідних для проведення дослідження бібліотек.

```

import warnings
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

from sklearn import linear_model

from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from xgboost import XGBClassifier

from sklearn import tree
from sklearn.metrics import accuracy_score
from sklearn.metrics import f1_score, recall_score,
precision_score, plot_roc_curve, roc_curve, roc_auc_score
from sklearn.model_selection import cross_val_predict
from sklearn.model_selection import cross_val_score

from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split

from sklearn.feature_selection import SelectPercentile
from sklearn.feature_selection import chi2 , f_classif

from sklearn.metrics import
mean_absolute_error,r2_score,mean_squared_error
from sklearn.feature_selection import RFE
warnings.filterwarnings('ignore')
data = pd.read_csv ('../input/diabetes-health-indicators-
dataset/diabetes_012_health_indicators_BRFSS2015.csv')

```

Рис. 3.2. Лістинг імпорту необхідних бібліотек

Далі потрібно відобразити фактори, які впливають на розвиток діабету і є наявними у дата сеті. Для цього використовуються наступні рядки коду:

```

data.head()
data.info()

```

У результаті отримаємо результат, який показує структуру дата сету у вигляді таблиці (рис 3.3) і тип характеристик осіб щодо їх схильності до захворювання цукровим діабетом (рис. 3.4)

	Diabetes_012	HighBP	HighChol	CholCheck	BMI	Smoker	Stroke	HeartDiseaseorAttack	PhysActivity	Fruits	...	A
0	0.0	1.0	1.0	1.0	40.0	1.0	0.0	0.0	0.0	0.0	...	1
1	0.0	0.0	0.0	0.0	25.0	1.0	0.0	0.0	1.0	0.0	...	0
2	0.0	1.0	1.0	1.0	28.0	0.0	0.0	0.0	0.0	1.0	...	1
3	0.0	1.0	0.0	1.0	27.0	0.0	0.0	0.0	1.0	1.0	...	1
4	0.0	1.0	1.0	1.0	24.0	0.0	0.0	0.0	1.0	1.0	...	1

Рис. 3.3. Представлення характеристик у вигляді таблиці

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 253680 entries, 0 to 253679
Data columns (total 22 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Diabetes_012                          253680 non-null float64
1   HighBP                                253680 non-null float64
2   HighChol                              253680 non-null float64
3   CholCheck                             253680 non-null float64
4   BMI                                    253680 non-null float64
5   Smoker                                 253680 non-null float64
6   Stroke                                 253680 non-null float64
7   HeartDiseaseorAttack                  253680 non-null float64
8   PhysActivity                          253680 non-null float64
9   Fruits                                253680 non-null float64
10  Veggies                               253680 non-null float64
11  HvyAlcoholConsump                     253680 non-null float64
12  AnyHealthcare                         253680 non-null float64
13  NoDocbcCost                           253680 non-null float64
14  GenHlth                               253680 non-null float64
15  MentHlth                              253680 non-null float64
16  PhysHlth                              253680 non-null float64
17  DiffWalk                              253680 non-null float64
18  Sex                                    253680 non-null float64
19  Age                                    253680 non-null float64
20  Education                             253680 non-null float64
21  Income                                253680 non-null float64
dtypes: float64(22)
memory usage: 42.6 MB

```

Рис. 3.4. Типи характеристик людей та їх схильності до захворювання цукровим діабетом

До характеристик, які впливають на розвиток цукрового діабету у даному дата сеті визначено наступні:

- високий артеріальний тиск;
- високий рівень холестерину;
- вживання препаратів для нормування рівня холестерину;
- індекс маси тіла;
- куріння;
- інсульт;
- серцеві напади;
- фізична активність;
- вживання овочів;
- вживання фруктів;
- вживання алкоголю;
- будь-які заходи піклування про здоров'я;
- загальний стан фізичного здоров'я;
- загальний стан психічного здоров'я;
- заняття спортивною ходьбою;
- вік;
- стать;
- освіта;
- дохід та ін.

Для встановлення кореляції між характеристиками дата сету запропоновано побудувати теплову карту шляхом виконання наступного програмного коду

```
plt.figure(figsize=(20,10))  
sns.heatmap(data.corr(), annot=True, cmap="YlGnBu")
```

Результат аналізу кореляцій представлено на рис. 3.5.

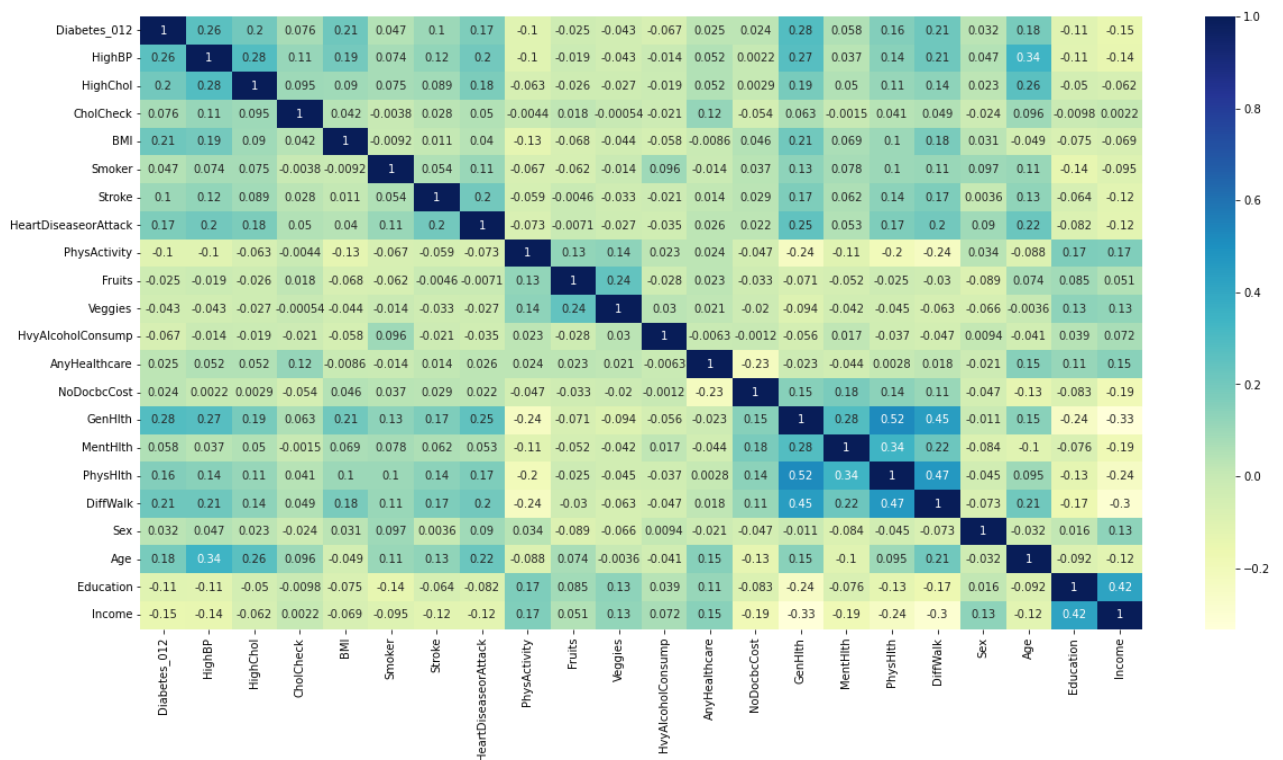


Рис. 3.5. Кореляції між даними у дата сеті

Після проведення трансформації даних (рис. 3.6), одержано кількість унікальних записів по кожному стовпцю дата сету, який інтерпретує кожную окрему характеристику.

```

#transform data
df_vis.Diabetes_012[df_vis['Diabetes_012'] == 0] = 'No Diabetes'
df_vis.Diabetes_012[df_vis['Diabetes_012'] == 1] = 'Pre Diabetes'
df_vis.Diabetes_012[df_vis['Diabetes_012'] == 2] = 'Diabetes'
df_vis.HighBP[df_vis['HighBP'] == 0] = 'No High'
df_vis.HighBP[df_vis['HighBP'] == 1] = 'High BP'
df_vis.HighChol[df_vis['HighChol'] == 0] = 'No High Cholesterol'
.....
df_vis.Smoker[df_vis['Smoker'] == 0] = 'No'
df_vis.Smoker[df_vis['Smoker'] == 1] = 'Yes'
df_vis.Stroke[df_vis['Stroke'] == 0] = 'No'
df_vis.Stroke[df_vis['Stroke'] == 1] = 'Yes'
df_vis.HeartDiseaseorAttack[df_vis['HeartDiseaseorAttack'] == 0] = 'No'
.....
unique_values = {}
for col in df_vis.columns:
    unique_values[col] = df_vis[col].value_counts().shape[0]
pd.DataFrame(unique_values, index=['unique value count']).transpose()

```

Рис. 3.6. Перетворення даних та визначення унікальних значень дата сету

	unique value count
Diabetes_012	3
HighBP	2
HighChol	2
CholCheck	2
BMI	84
Smoker	2
Stroke	2
HeartDiseaseorAttack	2
PhysActivity	2
Fruits	2
Veggies	2
HvyAlcoholConsump	2
AnyHealthcare	2
NoDocbcCost	2
GenHlth	5
MentHlth	31
PhysHlth	31
DiffWalk	2
Sex	2
Age	13
Education	6
Income	3

Рис. 3.7. Кількість унікальних записів у дата сеті

Розподіл даних у дата сеті за категоріями 0, 1, 2 має вигляд, як показано на рис. 3.8.

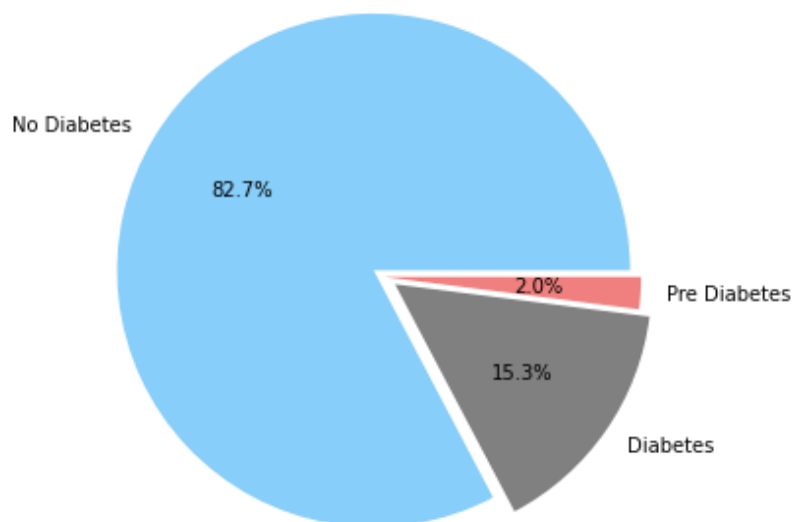


Рис. 3.8. Розподіл даних за цільовою змінною

Як видно з рис. 3.8 дані є дуже незбалансованими, тому необхідно проводити заходи балансування. Це дозволить в перспективі побудувати ефективні моделі прогнозування та виділення факторів, які найбільше впливають на розвиток діабету.

3.2. Виявлення кореляції між факторами впливу на розвиток діабету

Для визначення кореляції між факторами, які впливають на розвиток цукрового діабету потрібно програмно реалізувати візуалізацію залежності рівня цукру в крові від характеристик, наявних в дата сеті (рис. 3.9)

```
data.drop('Diabetes_012', axis=1).corrwith(data.Diabetes_012).plot(kind='bar', grid=True, figsize=(15, 6), title="Correlation with Diabetes_binary", color="blue");
```

Рис. 3.9. Фрагмент програмної реалізації виявлення кореляції

На рис. 3.10 показано залежність, яка реалізована програмним кодом, представленим на рис. 3.9.

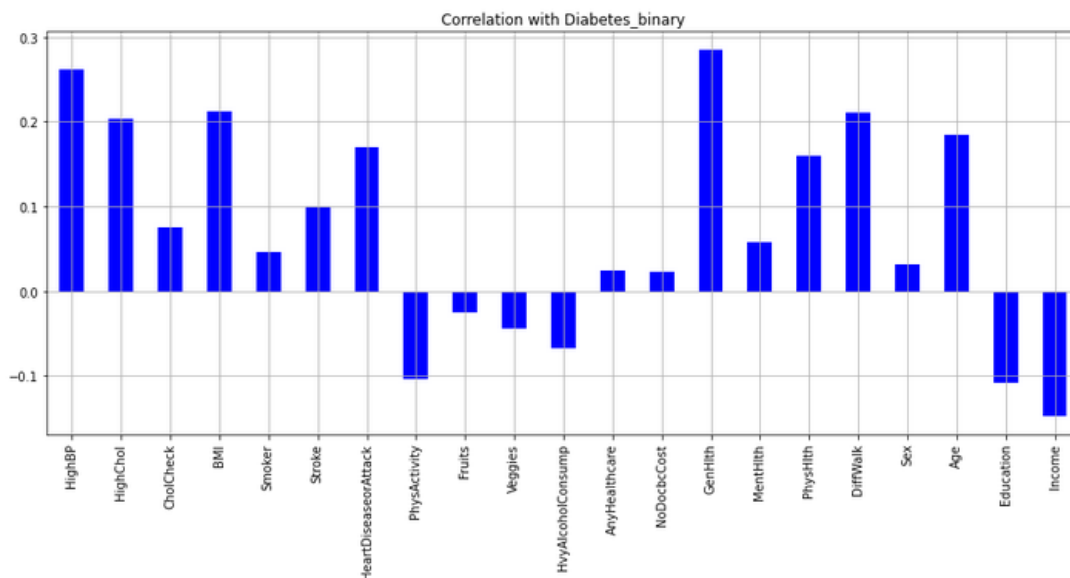


Рис. 3.10. Кореляція між факторами впливу на розвиток цукрового діабету та рівнем цукру в крові

Далі було проведено візуальний аналіз графіків щодо впливу окремо взятих факторів, які впливають на розвиток цукрового діабету. Так на рис. 3.11 і 3.12, в якості прикладів наведено відповідно розподіл хворих за статтю та тих, хто палить і не палить.

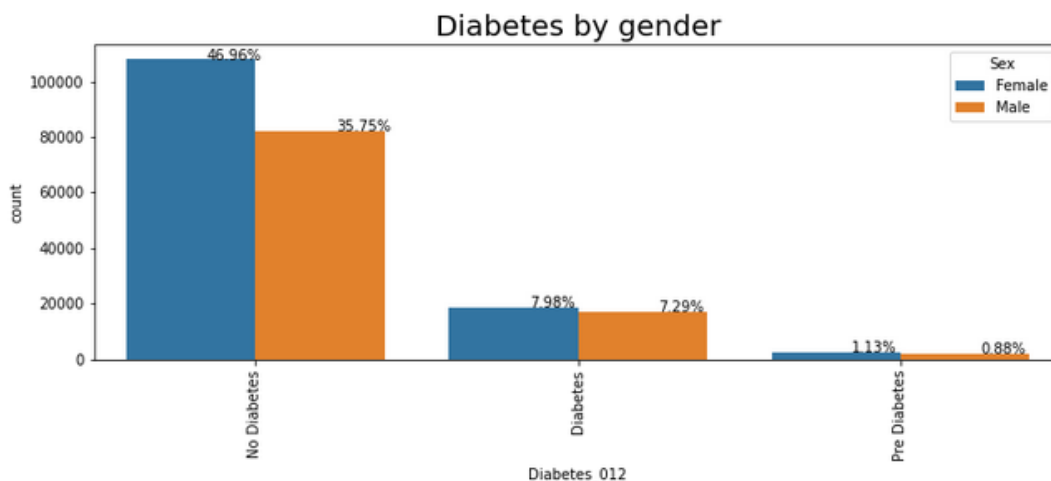


Рис. 3.11. Розподіл даних щодо захворюваність діабетом за статтю

Аналізуючи гістограму рис. 3.11 можна сказати, що серед наявних даних більше осіб чоловічої статі, які не хворіють діабетом, серед хворих з підвищеним рівнем глюкози трохи менше жінок і переддіабетний стан також практично рівномірно розподілений між жінками і чоловіками.

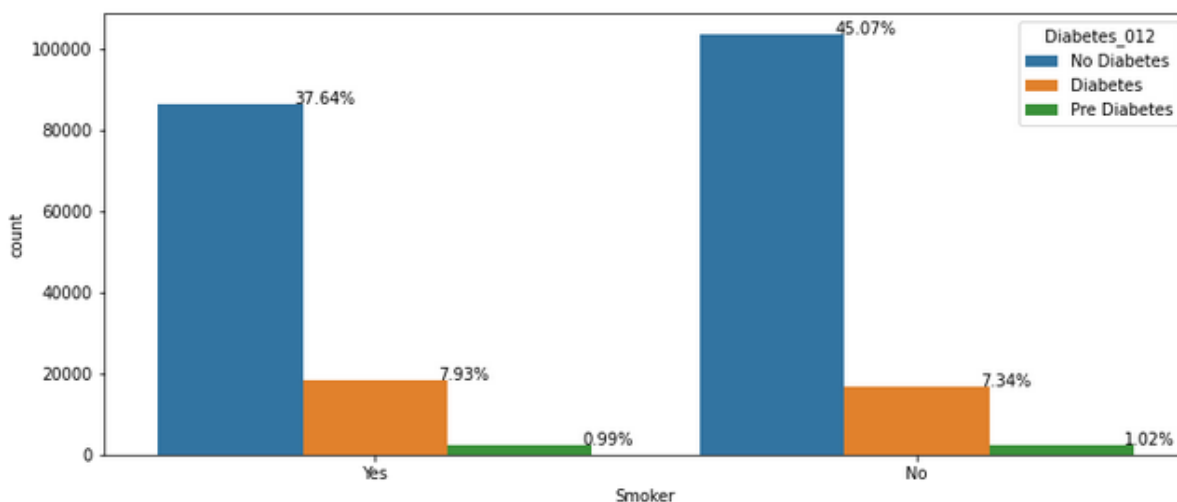


Рис. 3.12. Розподіл хворих за фактором куріння

Як видно з рис. 3.12, відсоток осіб, які не палять є вищим за відсоток тих, , хто не палять. Однак наявність ознак діабету практично однакова як для однієї групи, так і для іншої. Провівши аналіз для кожної з двадцяти однієї характеристики дата сету та їх комбінацій було встановлено, що існує залежність розвитку цукрового діабету для груп людей, які вживають алкоголь та палять. На рис. 3.13 показано графічно даний вплив.

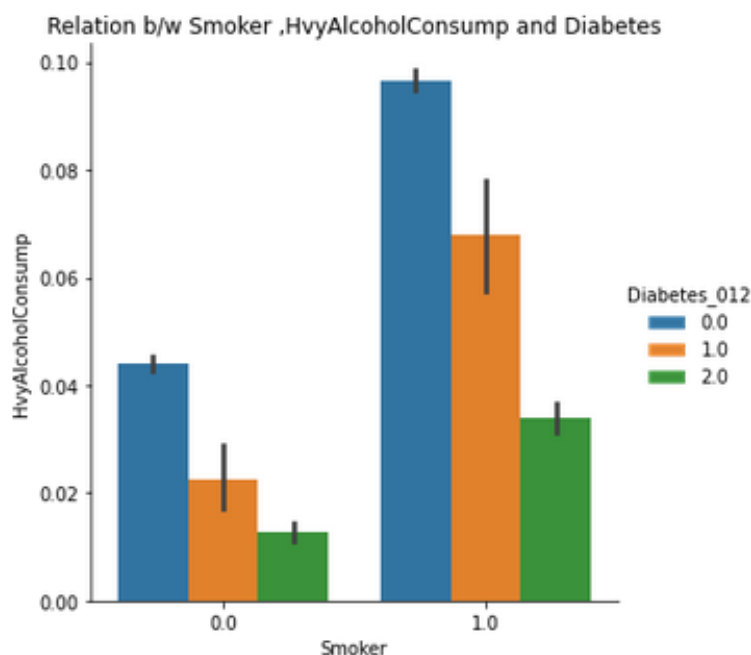


Рис. 3.13. Залежність розвитку діабету від куріння та вживання алкоголю

Як висновок з рис. 3.13 – алкоголь та куріння сприяють розвитку переддіабетного стану.

Високий рівень холестерину та високий тиск крові тісно пов'язані один з одним (рис. 3.14), оскільки люди з високим рівнем холестерину, як правило, мають високий тиск крові. Зв'язок між високим кров'яним тиском і високим рівнем холестерину існує в обох напрямках. Коли організм не може вивести холестерин із крові, надлишок холестерину може осідати на стінках артерій. Коли артерії стають жорсткими та звужуються через відкладення, серцю доводиться працювати понаднормово, щоб прокачувати через них кров. Це призводить до підвищення артеріального тиску.

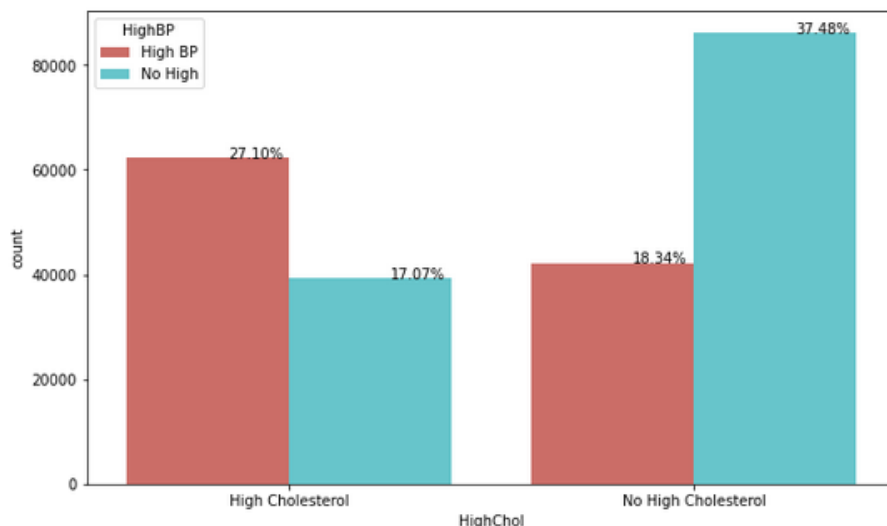


Рис. 3.14. Залежність рівня глюкози в крові від рівня холестерину

Підсумовуючи аналіз візуальної інформації щодо кореляції факторів впливу на розвиток цукрового діабету, можна зробити наступний висновок:

- чоловіки і жінки однаково вразливі до діабету;
- люди старше 45 років більш вразливі до діабету, ніж молодші;
- з віком кількість діабетиків також зростає;
- більше половини діабетиків страждають ожирінням, майже половина хворих на переддіабетний стан, страждають ожирінням відсоток діабетиків і преддіабетиків, які страждають від ожиріння та надмірної ваги, набагато вищий, ніж відсоток недіабетиків, які страждають від ожиріння та надмірної ваги
- коли освіта зростає, кількість діабетиків зменшується.
- люди з нижчими доходами мають більший ризик розвитку діабету, ніж люди з вищими доходами.
- генетика має великий вплив на діабет, тобто коли генетичні показники погані – ризик діабету швидко зростає;
- ментальне (психологічне) є основним фактором, який викликає діабет та залежить від його стабільності протягом тривалого часу;
- фізична активність знижує ризик діабету;
- вживання хоча б одного фрукта на день знижує ризик діабету;

– вживання хоча б одного овоча на день також знижує ризик діабету.

Провівши попередній аналіз даних, наявних в дата сеті потрібно провести більш детальний їхній препроцесинг.

3.3. Препроцесинг даних при прогнозуванні розвитку цукрового діабету

Препроцесинг даних варто почати з виявлення викидів (outlier). Для цього спочатку візуально потрібно переконатися у їх наявності. Програмний код, показаний у вигляді лістингу на рисунку 3.15 дозволяє згенерувати графік, представлений на рис. 3.16 для аналізу викидів.

```
plt.figure(figsize = (25,8))
u = sns.boxplot(palette = 'cool', data=data)
u.set_xticklabels(u.get_xticklabels(),rotation=45)
```

Рис. 3.15. Програмний код для візуалізації викидів

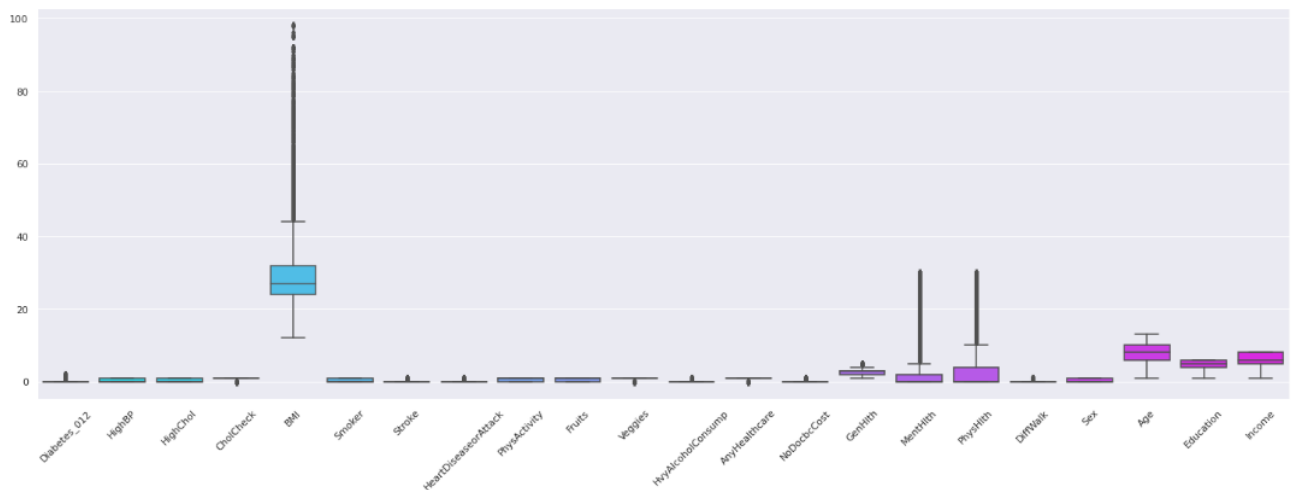


Рис. 3.16. Візуальне представлення викидів

Як видно з рис. 3.16 явними викидами дата сету є значення за ознакою індексу маси тіла (BMI), однак для того, щоб переконатися в істинності цього припущення, потрібно візуалізувати більш детально значення за всіма ознаками (рис. 3.17).

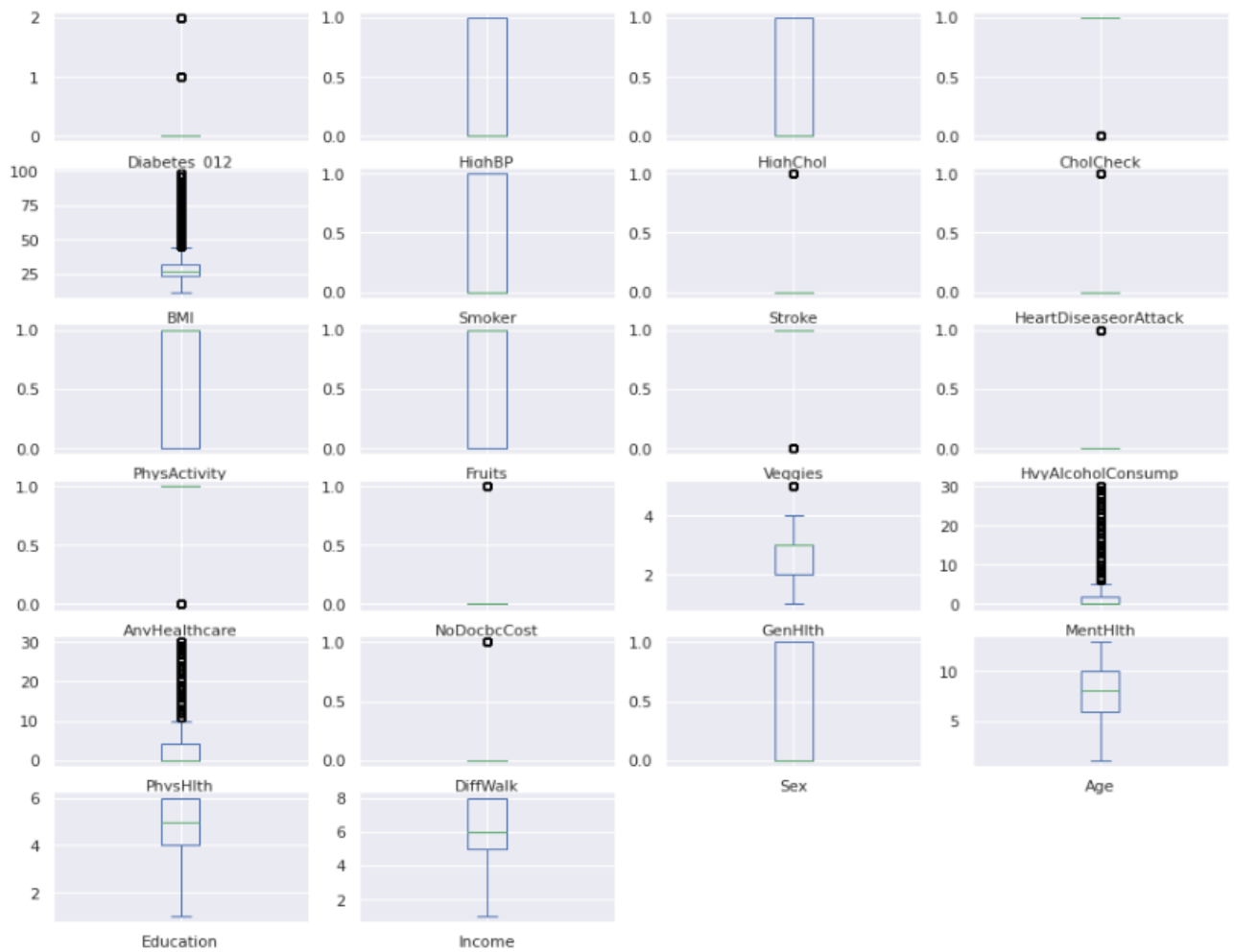


Рис. 3.17. Розкид значень за усіма ознаками дата сету при дослідженні розвитку цукрового діабету

Знаходження конкретних значень викидів за ознакою надлишкової маси тіла передбачає виконання програмного коду, наведеного на рис. 3.18.

```
plt.figure(figsize = (12,6))
plt.subplot(1, 2, 1)
sns.boxplot(data=data,y='BMI',color='#cc6699')
plt.subplot(1, 2, 2)
sns.scatterplot(data=data,x='Diabetes_012',y='BMI',color='#cc6699')
plt.show()
```

Рис. 3.18. Програмний код виявлення викидів за ознакою індексу маси тіла

Якщо візуально представити значення за ВМІ, то розподіл за значенням >70 і ≤ 70 матиме вигляд, як показано на рис. 3.19.

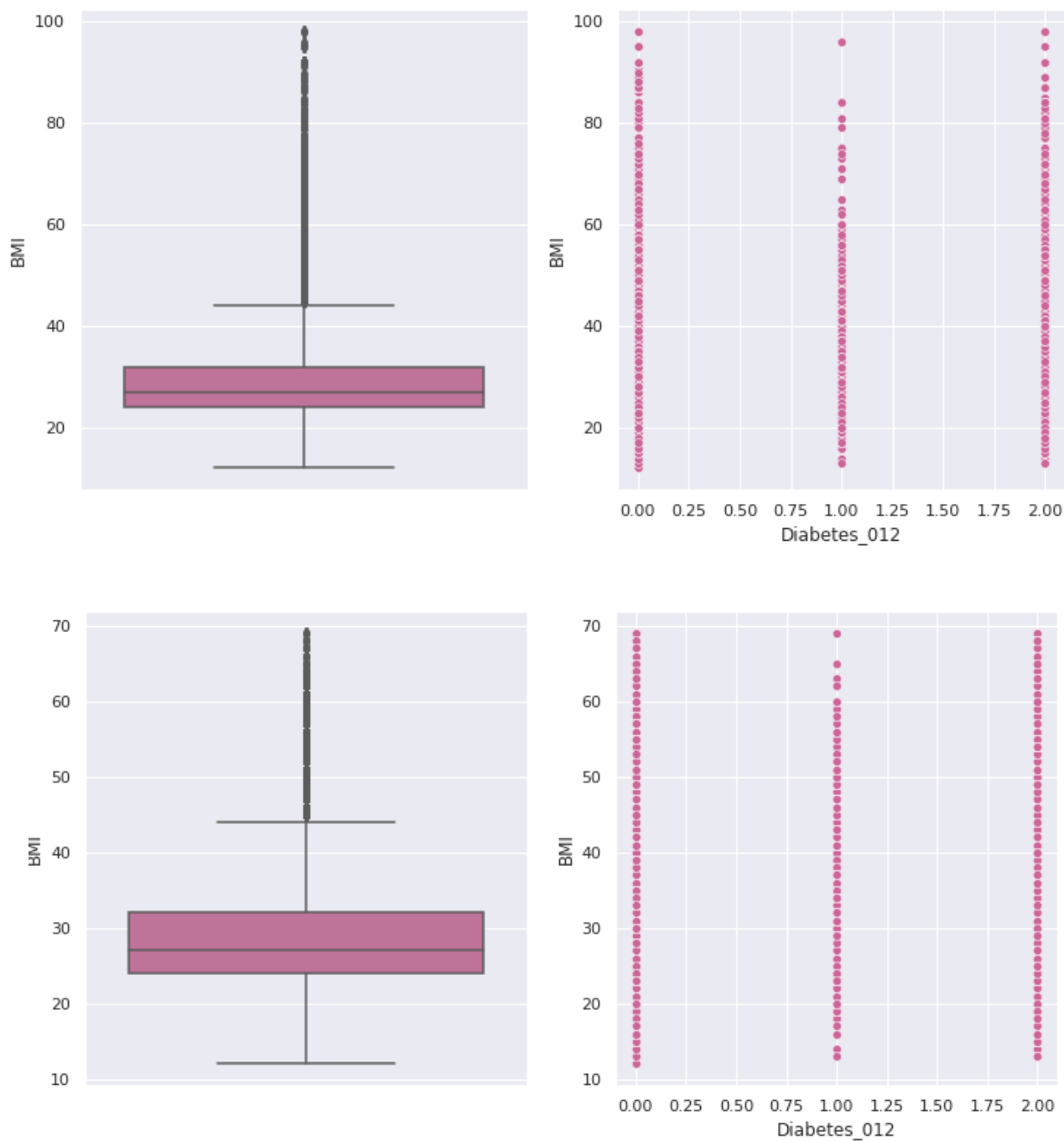


Рис. 3.19. Виявлення викидів за ознакою індексу маси тіла

Після виконання таких маніпуляцій потрібно видалити дані, які є викидами і перестворити дата сет. Для цього застосовується лістинг програмного коду, показаний на рис. 3.20.

```

y = df['Diabetes_012']
x = df.drop(['Diabetes_012'], axis=1)
x_train , x_test , y_train , y_test = train_test_split(x,y ,
test_size= 0.25 , random_state=42)
from imblearn.over_sampling import SMOTE
sm = SMOTE(random_state=42)
X_res, y_res = sm.fit_resample(x_train, y_train)

```

Рис. 3.20. Видалення викидів

Маючи оновлений дата сет, потрібно перейти до моделювання алгоритмів прогнозування розвитку цукрового діабету та визначення найбільш важливих факторів, які на нього впливають.

3.4. Реалізація алгоритмів прогнозування розвитку захворюваності цукровим діабетом

Для вибору та проведення експериментів щодо прогнозування рівня глюкози в крові для трьох категорій людей: ті, у яких нормальний рівень цукру в крові, ті, хто хворіє діабетом першого типу, ті, хто хворіє діабетом другого типу, потрібно застосувати методи машинного навчання, оскільки кількість даних є надзвичайно велика, а визначити і врахувати кореляції між факторами, або їх групами вручну практично неможливо.

Потрібно відмітити, що при прогнозуванні рівня цукру в крові необхідно розв'язати задачу класифікації та визначити сукупність і пріоритет факторів, які найбільше впливають на розвитку цукрового діабету. У даному випадку, прогнозування трактується як prediction, а не прогнозування розвитку цукрового діабету в часі – forecast. Як приклад задач, методів і алгоритмів, для яких можна застосовувати інструменти машинного навчання належать такі, які проілюстровані на рис. 3.21

Як було зазначено раніше, прогнозування розвитку цукрового діабету буде вирішуватися за допомогою методів класифікації.

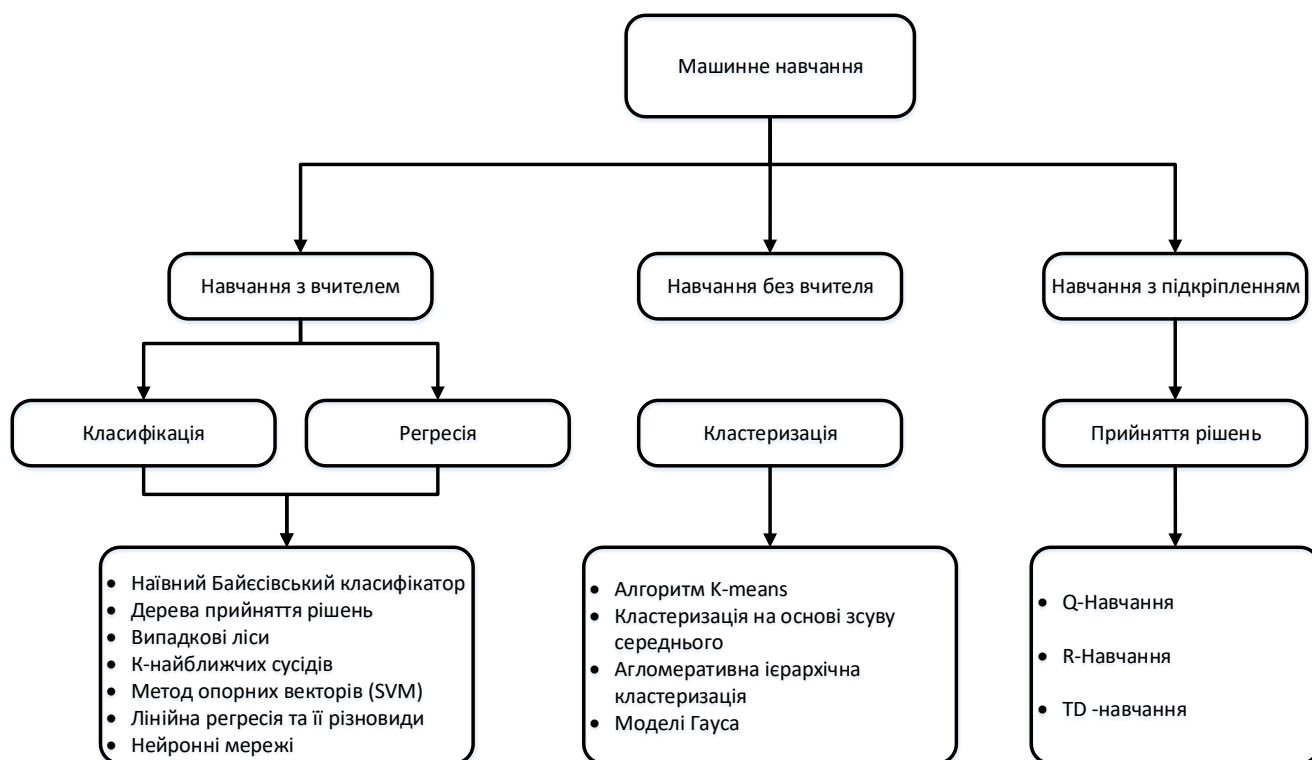


Рис. 3.21. Застосування методів і алгоритмів машинного навчання

З практичної точки зору, для розв'язку задачі такого класу найбільш ефективними є алгоритми:

- логістичної регресії;
- дерев прийняття рішення;
- XGBoost;
- «випадкового лісу».

Логістична регресія, як і проста лінійна регресія були запозичені з розділу математичної статистики. Відмінною рисою логістичної регресії є те, що значення функції представляється імовірністю.

На вхід логістичної регресії, в принципі як і лінійної, подають одну або декілька незалежних змінних та проводять обчислення впливу чи залежності від цільової змінної. Відмінність між логістичною та лінійною регресією полягає у застосуванні сигмоподібної функції, яка дає можливість прогнозувати неперервну змінну в інтервалі від 0 до 1 за будь-яких значень незалежних змінних. Фактично, логістична регресія представляє собою розподіл Бернуллі. Приклад лінійної і логістичної регресій показано на рис. 3.22.

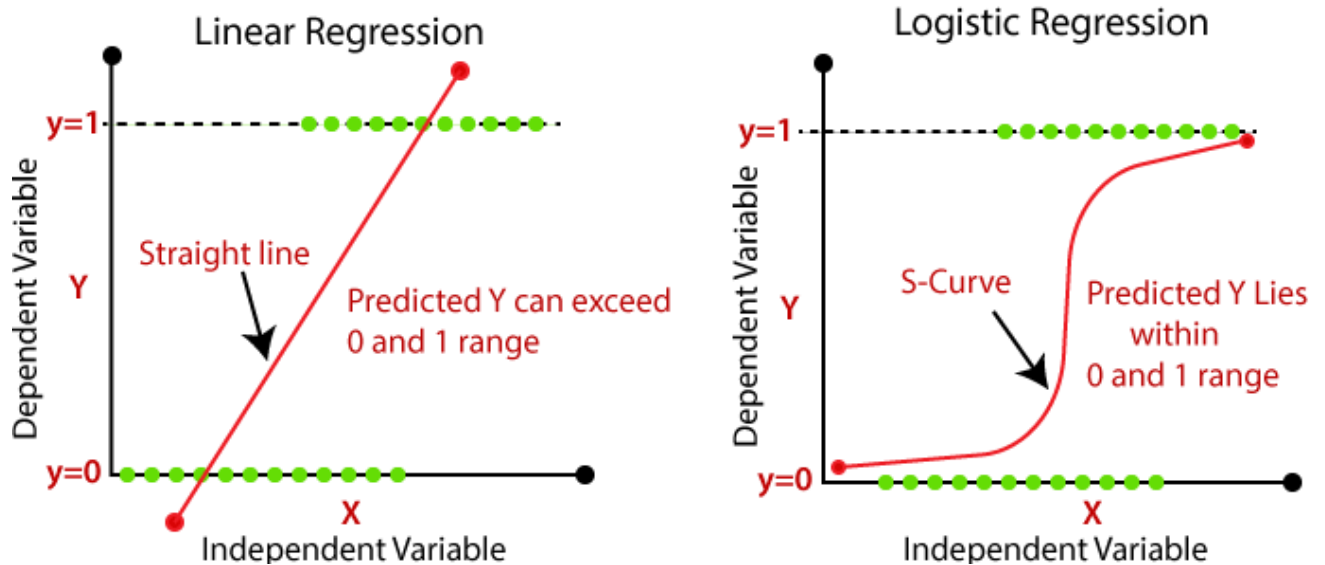


Рис. 3.22. Приклад лінійної і логістичної регресій

Лістинг програмного коду застосування логістичної регресії показаний на рис. 3.23.

```
lg = LogisticRegression()
lg.fit(X_res, y_res)
y_pred_train_lg = lg.predict(X_res)
acc_train_lg = accuracy_score(y_res, y_pred_train_lg)

y_pred_test_lg = lg.predict(x_test)
acc_test_lg = accuracy_score(y_test, y_pred_test_lg)
print(acc_train_lg)
print(acc_test_lg)
print(classification_report(y_test, y_pred_test_lg))
```

Рис. 3.23. Реалізація логістичної регресії

Оцінка за критеріями щодо точності прогнозування з використанням підходу логістичної регресії показано на рис. 3.24.

	precision	recall	f1-score	support
0.0	0.93	0.61	0.74	47315
1.0	0.03	0.32	0.06	1125
2.0	0.34	0.56	0.42	8856
accuracy			0.60	57296
macro avg	0.43	0.50	0.41	57296
weighted avg	0.82	0.60	0.68	57296

Рис. 3.24. Оцінка якості роботи алгоритму логістичної регресії

Реалізація оцінки точності прогнозування за допомогою кривої ROC-AUC становить близько 73%, програмний код якої показано на рис. 3.25.

```

y_pred_prob_lg = lg.predict_proba(x_test)
roc_auc_score_lg= roc_auc_score(y_test, y_pred_prob_lg, multi_class="ovr")

print('ROC AUC Score: ',roc_auc_score_lg)

```

ROC AUC Score: 0.7332104791112362

Рис. 3.25. Оцінка точності прогнозування на основі ROC-AUC

Матриця помилок показує відповідність щодо кількості правильно спрогнозованих значень і наведена на рис. 3.26.

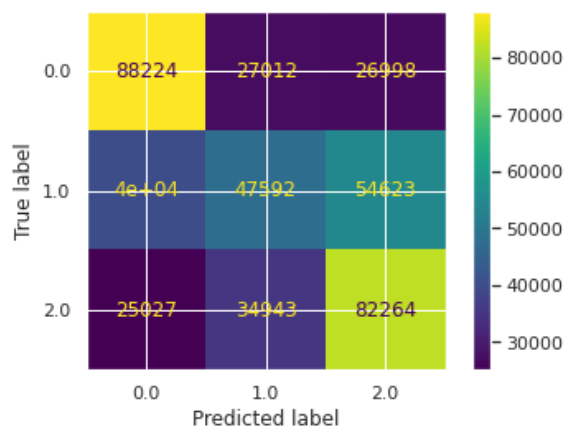


Рис. 3.26. Матриця помилок при прогнозуванні розвитку діабету

Наступний алгоритм прогнозування розвитку цукрового діабету полягає в реалізації підходу випадкового лісу. Для цього потрібно виконати програмний код представлений на рис. 3.27.

```
### 2) Random Forest Classification

from sklearn.ensemble import RandomForestClassifier
rf = RandomForestClassifier(n_estimators =
100,max_depth=16,max_features=10)
rf.fit(X_res, y_res)
y_pred_train_rf = rf.predict(X_res)
acc_train_rf = accuracy_score(y_res, y_pred_train_rf)

y_pred_test_rf = rf.predict(x_test)
acc_test_rf = accuracy_score(y_test, y_pred_test_rf)
print(acc_train_rf)
print(acc_test_rf)
```

Рис. 3.27. Застосування класифікатора випадкового лісу

Якість результатів класифікації при виконанні, наведеного вище коду показано на рис. 3.28.

	precision	recall	f1-score	support
0.0	0.87	0.91	0.89	47315
1.0	0.06	0.00	0.01	1125
2.0	0.44	0.38	0.41	8856
accuracy			0.81	57296
macro avg	0.46	0.43	0.44	57296
weighted avg	0.79	0.81	0.80	57296

Рис. 3.28. Оцінка якості результатів класифікації з використанням підходу випадкового лісу

Значення оцінки точності прогнозування з використанням метрики ROC AUC становить близько 75% та проілюстровано на рис. 3.29.

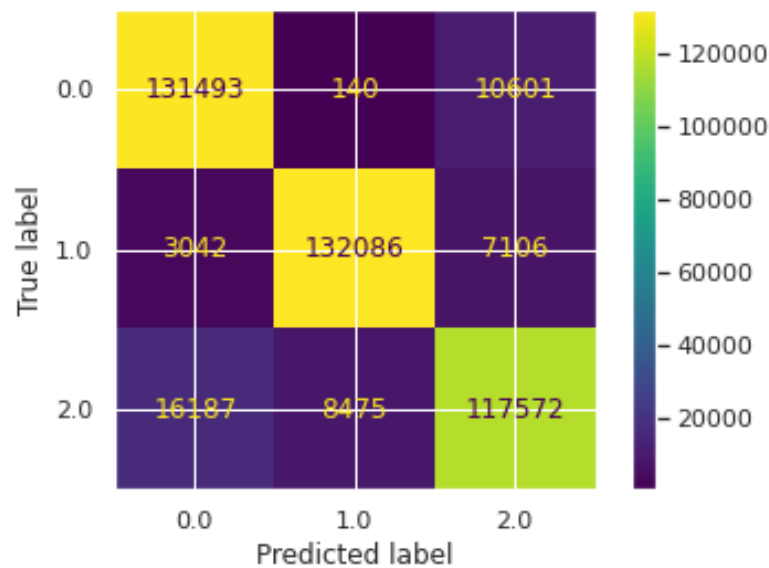


Рис. 3.29. Оцінка прогнозування за допомогою метрики ROC AUC

Реалізація алгоритму XGBoost для прогнозування рівня цукру в крові показана у лістингу на рис. 3.30.

```
xgb= XGBClassifier(max_depth=10)
xgb.fit(X_res, y_res)
y_pred_train_xgb = xgb.predict(X_res)
acc_train_xgb = accuracy_score(y_res, y_pred_train_xgb)

y_pred_test_xgb = xgb.predict(x_test)
acc_test_xgb = accuracy_score(y_test, y_pred_test_xgb)
print(acc_train_xgb)
print(acc_test_xgb)
```

Рис. 3.30. Лістинг реалізації алгоритму XGBoost

При оцінюванні якості роботи алгоритму XGBoost досягнуто точності класифікації на різних метриках на рівні від 85% до 92%, що показано на рис. 3.31.

	precision	recall	f1-score	support
0.0	0.85	0.96	0.90	47315
1.0	0.20	0.00	0.00	1125
2.0	0.49	0.22	0.30	8856
accuracy			0.83	57296
macro avg	0.51	0.39	0.40	57296
weighted avg	0.78	0.83	0.79	57296

Рис. 3.31. Результати оцінювання алгоритму XGBoost

Значення метрики ROC AUC при використанні алгоритму XGBoost становить близько 73% та проілюстровано на рис. 3.32.

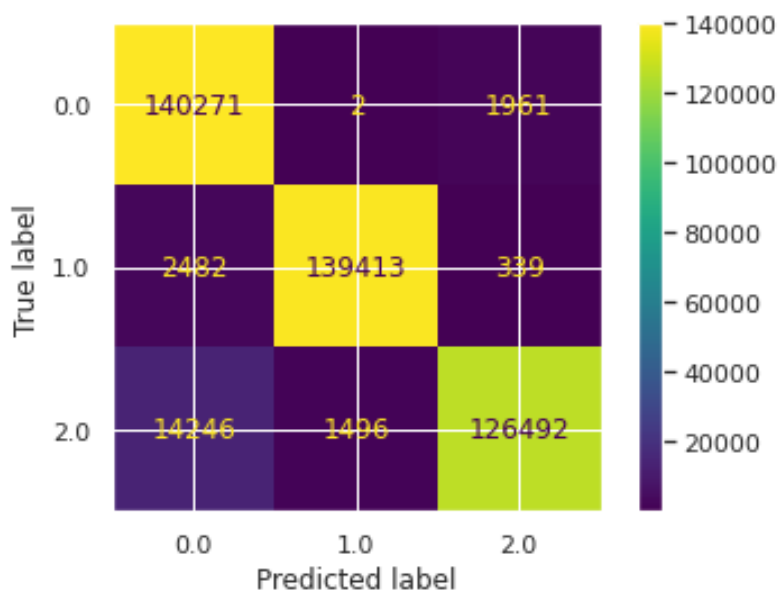


Рис. 3.32. Оцінка точності прогнозування за допомогою ROC AUC

Ще один алгоритм, який може забезпечити високу точність прогнозування розвитку цукрового діабету – дерева прийняття рішень. На рис. 3.33 показано програмний код реалізації класифікації за допомогою дерев прийняття рішень.

```

dt= DecisionTreeClassifier(max_features=10 , max_depth=16)
dt.fit(X_res, y_res)
y_pred_train_dt = dt.predict(X_res)
acc_train_dt = accuracy_score(y_res, y_pred_train_dt)

y_pred_test_dt = dt.predict(x_test)
acc_test_dt = accuracy_score(y_test, y_pred_test_dt)
print(acc_train_dt)
print(acc_test_dt)

```

Рис. 3.33. Класифікатор дерева прийняття рішень

Результати якості роботи класифікатора дерев прийняття рішень показана на рис. 3.34.

	precision	recall	f1-score	support
0.0	0.87	0.89	0.88	47315
1.0	0.04	0.02	0.03	1125
2.0	0.38	0.35	0.37	8856
accuracy			0.79	57296
macro avg	0.43	0.42	0.42	57296
weighted avg	0.78	0.79	0.78	57296

Рис. 3.34. Результати якості прогнозування за допомогою дерев прийняття рішень

Оцінка якості за допомогою метрики ROC AUC алгоритму дерев прийняття рішень становить приблизно 70%.

Порівняння точності класифікації і прогнозування розвитку цукрового діабету виконуються за допомогою програмного коду представленого на рис. 3.35.

```

Performance = pd.DataFrame(
    data = {
        'Model_after_resamplling': ['LogReg', 'RF', 'XGB', 'DT'],
        'Test_score': [accuracy_score(y_test, y_pred_test_lg),
                      accuracy_score(y_test, y_pred_test_rf),
                      accuracy_score(y_test, y_pred_test_xgb),
                      accuracy_score(y_test, y_pred_test_dt)],

        'ROC_AUC_Score': [roc_auc_score_lg,
                          roc_auc_score_rf,
                          roc_auc_score_xgb,
                          roc_auc_score_dt]
    }
)

```

Рис. 3.35. Порівняння алгоритмів прогнозування

Результат порівняння якості алгоритмів при прогнозуванні розвитку цукрового діабету показаний на рис. 3.36.

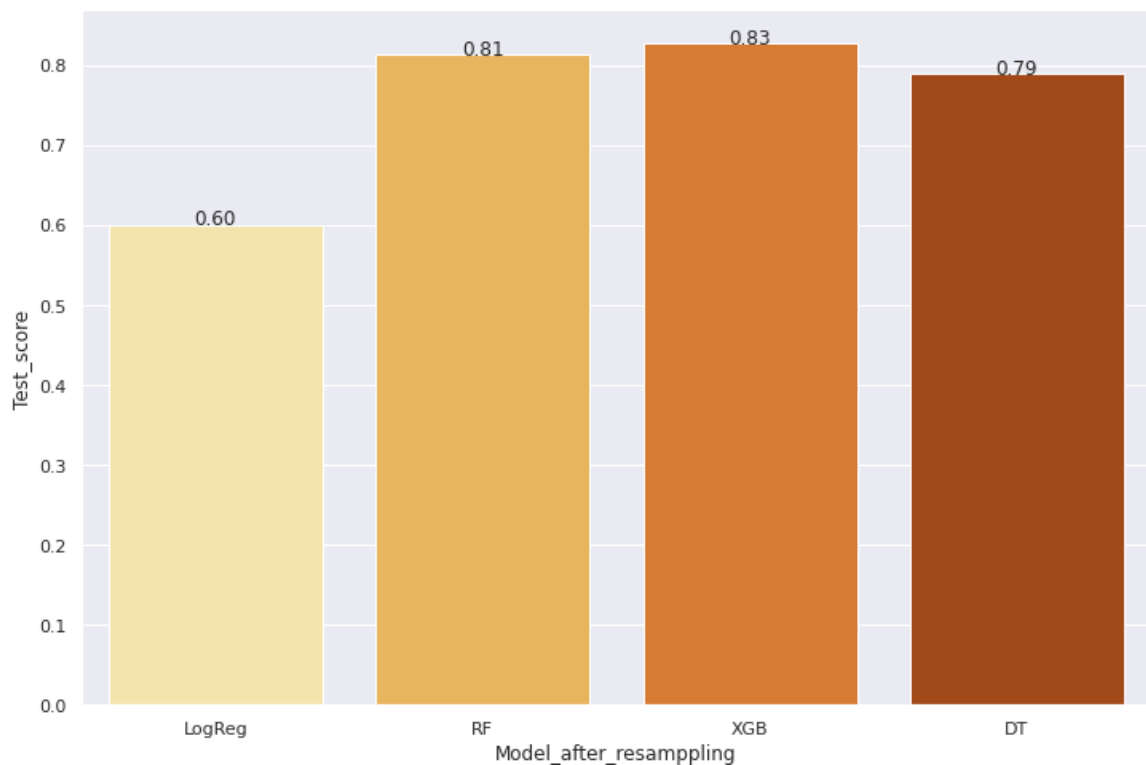


Рис. 3.36. Оцінка ефективності алгоритмів

Як видно з рис. 3.37 для підвищення точності прогнозування необхідно визначити найбільш важливі фактори впливу на розвиток цукрового діабету. Для

цього було реалізовано відповідні маніпуляції у результаті якого одержано ранжований список важливості ознак розвитку цукрового діабету (рис. 3.37).

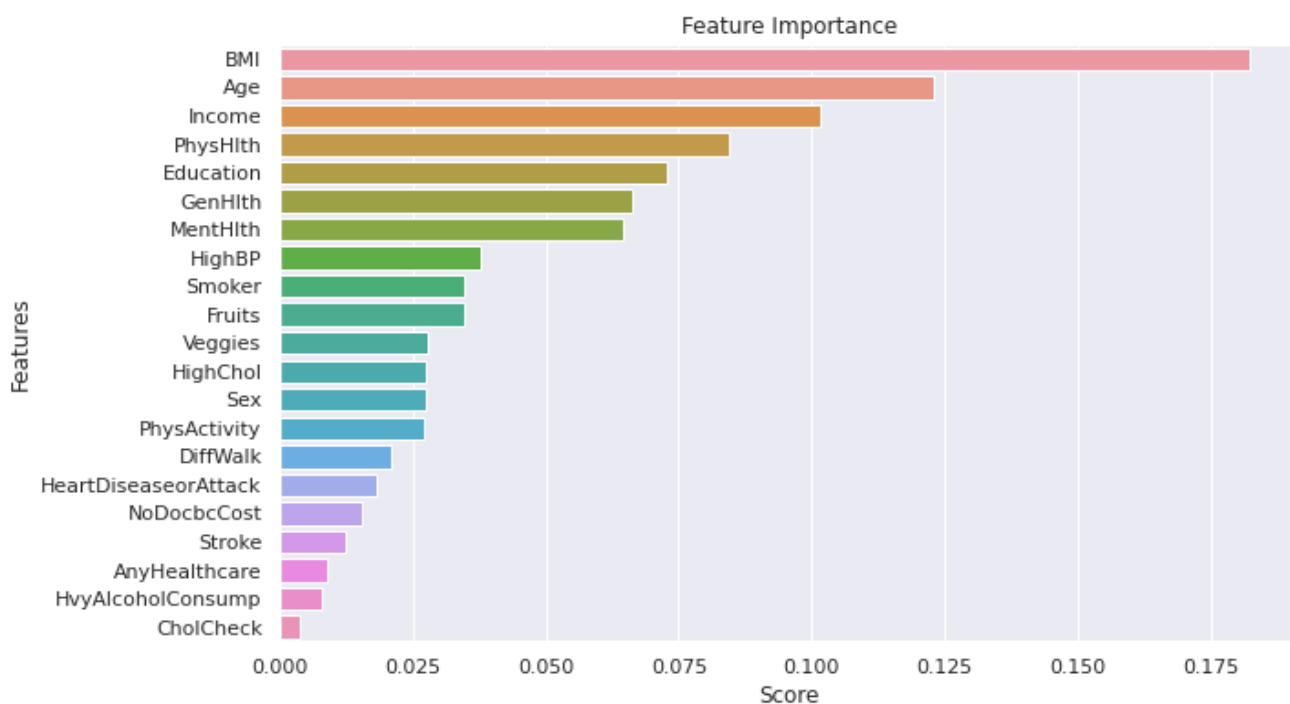


Рис. 3.37. Важливість ознак щодо впливу на розвиток цукрового діабету

Провівши повторні моделювання на основі 14 ознак з найбільшою вагою впливу на розвиток цукрового діабету досягнуто результатів точності прогнозування на рівні 99%.

Підсумовуючи результати проведеного моделювання одержано наступні результати:

1. Основними ознаками діабету є: високий кров'яний тиск, високий рівень холестерину, індекс маси тіла, інсульт, загальний рівень, розумовий рівень, фізичний рівень, вік, освіта та дохід.

2. Характерними ознаками, які в комплексі збільшують ризик діабету, є: паління та надмірне споживання алкоголю, інсульт та серцево-судинні хвороби чи інфаркти, високий артеріальний тиск і рівень холестерину.

3. Функціональними змінними, які є найменше впливають на розвиток діабету, але можуть допомогти зменшити такий ризик є: фізична активність,

вживання фруктів та овочів, регулярна перевірка рівня холестерину, турбота про здоров'я .

4. Через незбалансовані дані, базова оцінка точності надавала не високі значення, тому використано критерії для оцінки точності прогнозування, такі як точність/специфічність, запам'ятовування/чутливість, оцінка F1 та AUC.

5. Після попередньої обробки та повторної вибірки даних вони стали збалансованішими та прогнозування забезпечило більшу точність.

6. Існують чіткі відмінності в показниках між застосуванням алгоритмів після повторної вибірки даних.

7. XGBoost є найкращим алгоритмом, який має оцінку в тесті та оцінку AUC.

8. Після вибору найбільш важливих ознак розвитку діабету для прогнозування обрано 70% (14 ознак) і точність зросла до 98-100%.

3.5. Висновки до розділу

Основні результати даного розділу полягають в наступному:

1. Проведено аналіз даних, які включають 24 фактори впливу на появу і розвиток цукрового діабету, обґрунтовано використання мови програмування Python та відповідних відкритих бібліотек, які в подальшому забезпечать моделювання процесу опрацювання даних, визначення факторів з найбільшою вагою впливу на розвиток цукрового діабету та прогнозування схильності людей до цього захворювання .

2. Реалізовано процедури визначення кореляції між цільовою змінною та іншими незалежними змінними, наявними у даних, що дало змогу встановити, що найбільше появу і розвиток різного типу діабету провокує куріння, алкоголь, психологічний стан і доходи, які одержує людина, а в комплексі факторів – паління та надмірне споживання алкоголю, інсульт та серцево-судинні хвороби чи інфаркти, високий артеріальний тиск і рівень холестерину.

3. Реалізовано алгоритми прогнозування (класифікації) розвитку цукрового діабету на основі підходів дерев прийняття рішень, випадкового лісу, логістичної регресії та XGBoost і встановлено, що при врахуванні усіх факторів впливу на розвиток цукрового діабету найбільшу точність прогнозування забезпечує XGBoost на рівні 83%. Даний алгоритм застосовано після вибору 14 найбільш впливових факторів і при цьому точність підвищилась до інтервалу 98-100%.

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

Метою кваліфікаційної роботи магістра є дослідження методів та програмно-апаратних засобів опрацювання даних у системах моніторингу рівня цукру в крові. Оскільки, проведення робіт зі створення та аналізу математичного і програмно-апаратного забезпечення передбачає використання комп'ютерної техніки, зокрема ПК та периферійних пристроїв, то обов'язком виконавця такого процесу є забезпечення оптимальних умов праці з охорони праці і техніки безпеки.

Для забезпечення, контролю та моніторингу за дотриманням виконання вимог з охорони праці керівництво компаній, які планують використовувати програмно-апаратні засоби опрацювання даних у системах моніторингу рівня цукру в крові, повинно організувати функціонування системи управління охороною праці, а саме:

- створити відповідні служби і призначити посадових осіб, які забезпечують вирішення конкретних питань охорони праці, затвердити інструкції про їхні обов'язки, права та відповідальність за виконання покладених на них функцій, а також контроль за їх додержанням;
- розробити за участю сторін колективного договору і реалізувати комплексні заходи для досягнення встановлених нормативів та підвищення існуючого рівня охорони праці;
- забезпечити виконання необхідних профілактичних заходів відповідно до обставин, що змінюються;
- впроваджувати прогресивні технології, досягнення науки і техніки, засоби механізації та автоматизації виробництва, вимоги ергономіки, позитивний досвід з охорони праці тощо;
- забезпечувати належне утримання будівель та споруд, виробничого обладнання та устаткування, моніторинг за їх технічним станом, зокрема

комп'ютерної техніки, серверів та оргтехніки;

- забезпечувати усунення причин, що призводять до нещасних випадків, професійних захворювань, та здійснення профілактичних заходів, визначених комісіями за підсумками розслідування цих причин;

- організовувати проведення аудиту охорони праці, лабораторних досліджень умов праці, оцінку технічного стану виробничого обладнання та устаткування, атестацій робочих місць на відповідність нормативно-правовим актам з охорони праці в порядку і строки, що визначаються законодавством, та за їх підсумками вживає заходів з усунення небезпечних і шкідливих для здоров'я виробничих факторів;

- здійснювати контроль за додержанням працівником технологічних процесів, правил поведінки з машинами, механізмами, устаткуванням та іншими засобами виробництва, використанням засобів колективного та індивідуального захисту, виконанням робіт відповідно до вимог з охорони праці.

На робочих місцях осіб, які проектують систему моніторингу та прогнозування рівня цукру в крові необхідно забезпечити дотримання вимог НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Для забезпечення оптимальних умов праці при проведенні досліджень кваліфікаційної роботи, необхідно передбачити також відповідність мікроклімату у приміщеннях. Категорія робіт при експлуатації системи моніторингу рівня цукру в крові належить до легкої – Іб.

Для того щоб визначити, чи відповідає повітряне середовище певного приміщення встановленим нормам, необхідно кількісно оцінити кожний з його параметрів. Оптимальні показники мікроклімату, які необхідно забезпечити у приміщеннях, де експлуатуються ПК у теплу пору року повинні становити: температура – 22-24 °С, відносна вологість – 40-60%, швидкість руху повітря 0,1 м/с.

Окрім, забезпечення оптимальних показників мікроклімату, необхідно передбачити ще й оптимальні показники шуму та вібрації на робочих місцях.

Робоче місце працівників, які розробляють і досліджують програмно-апаратні засоби моніторингу і прогнозування рівня цукру в крові, можна прирівняти до робочих місць у приміщеннях конструкторських бюро, програмістів обчислювальних машин, лабораторій для теоретичних робіт і опрацювання експериментальних даних, прийому хворих в медпунктах і відповідно необхідно передбачити відповідні рівні звукового тиску.

Основними вимогами, визначеними у НПАОП 0.00-7.15-18 є:

- площу для одного робочого місця оператора визначають згідно норми не менше 6,0 кв.м, а об'єм - не менше 20,0 куб.м.
- заземлені конструкції, що знаходяться в приміщеннях, де розміщені робочі місця операторів (батареї опалення, водопровідні труби, кабелі із заземленим відкритим екраном), мають бути надійно захищені діелектричними щитками або сітками з метою недопущення потрапляння працівника під напругу.
- приміщення, де розміщені робочі місця операторів, крім приміщень, у яких розміщені робочі місця операторів великих ЕОМ загального призначення (сервер), повинні бути оснащені системою автоматичної пожежної сигналізації.

Для штучного освітлення нормованим параметром виступає $E_{\text{мін}}$ – мінімальний рівень освітленості, та $K_{\text{п}}$ – коефіцієнт пульсації світлового потоку. Для забезпечення оптимальних умов праці необхідно передбачити коефіцієнт пульсації світлового потоку на рівні не більшому, ніж 20% відповідно до ДБН В.1.1-7-2016. Оскільки, робота щодо дослідження системи моніторингу і прогнозування рівня цукру в крові належить до IV розряду зорових робіт, то мінімальний рівень штучного освітлення, який необхідно передбачити складає 300...500 Лк.

Сукупність заходів щодо забезпечення охорони праці, починаючи від виконання встановлених законами України норм та правил, а також виконання правил техніки безпеки, гарантує особам, які працюють над дослідженням системи моніторингу та прогнозування рівня цукру в крові, безпеку праці та нівелювання негативних факторів впливу на їх здоров'я.

4.2. Особливості роботи та розлади здоров'я користувачів комп'ютерів, що формується під впливом роботи за комп'ютером.

При роботі з комп'ютером протягом тривалого періоду часу, а також при недотриманні відповідних вимог охорони праці у користувачів спостерігаються розлади, які узагальнено назвою комп'ютерний зоровий синдром.

Комп'ютерний зоровий синдром (КЗС) – комплекс порушень здоров'я, який може виникати у користувачів персональних комп'ютерів (ПК) [26]. Діагноз ставлять, якщо людина, що працює за ПК протягом двох годин, висловлює хоча б дві з десяти скарг:

- головний біль;
- сльозотеча;
- різь;
- туман;
- двоїння;
- свербіж;
- важкість в очах;
- фотофобія;
- миготіння знаків на екрані;
- нудота.

У користувачів ПК дуже поширені кон'юнктивіти і блефарити, патогенетично пов'язані з КЗС.

Синдром розвивається при умові, що робоче місце організовано неправильно – у користувача незручне крісло, відсутні пюпітри для паперів, підставки для ніг та кистей рук, не встановлена висота і нахил монітора відносно очей, відстань від очей до екрана. За таких умов тіло людини при роботі займає вимушене положення: спина статично напружена, шия витягнута, плечі жорстко фіксовані. Напружені м'язи погіршують кровотік у сонних артеріях, а недостатнє кровозабезпечення головного мозку веде до очманіння, появи головного болю. На фоні шийного

остеохондрозу з'являється відчуття випирання очних яблук, туману в очах, мушок та райдужних кіл у полі зору. Розвитку КЗС сприяє поганий мікроклімат приміщення, значна загальна іонізація та мікробне забруднення, а також куріння.

Національною радою з наукових досліджень США для стану зорового дискомфорту був уведений термін "астенопія", який означає "будь-які суб'єктивні зорові симптоми чи емоційний дискомфорт, що є результатом зорової діяльності". Симптоми астенії були класифіковані на "очні" (біль, печія та різь в очах, почервоніння повік та очних яблук, ломота у надбрівній частині тощо) та "зорові" (пелена перед очима, мерехтіння, швидка втома під час зорової роботи та ін.).

У операторів ВДТ "очні" симптоми трапляються частіше, ніж "зорові", причому частота проявів астенії вища у жінок, ніж у чоловіків і більше виражена в осіб середнього і старшого віку. Причиною вважається електромагнітне випромінювання від ВДТ.

При роботі з ВДТ основне навантаження припадає на всі елементи зорового аналізатора.

Робота з ВДТ може призвести до розвитку короткозорості, так як у користувачів комп'ютерів, в основному, "працює" ближній зір.

При аналізі зорової роботи операторів ВДТ, встановлено, що через дві години частота флуктуацій акомодатції зменшується, а внесок низькочастотної компоненти підвищується. Це може бути причиною скарг на втому зорового аналізатора. Тривала робота на ВДТ може призвести до розвитку короткозорості, оскільки у користувачів ВДТ головним чином "працює" ближній зір.

У 100 пацієнтів із 150, які працювали на ВДТ по шість годин на день протягом чотирьох років, були виявлені проблеми з фокусуванням зору.

Робота за комп'ютером характеризується також тим, що постійний напружений погляд на екран монітора зменшує частоту моргання. При цьому погіршується зволоження поверхні очного яблука сльозовою рідиною, яка захищає рогівку ока від висихання, пилу та інших забруднень. Це може призвести до виникнення так званого синдрому Сікка: рогівка висихає і мутніє, і як наслідок розвивається сліпота.

Також при напруженій зоровій роботі за ЕОМ можуть бути не лише порушення функції зору, а й виникнення головного болю, посилення нервово-психічного напруження, зниження працездатності.

Виникнення та розвиток патології зорової функції зумовлені:

1. Умовами зорової роботи на ВДТ (зменшення вільного руху очей, зменшення функціонального поля сітківки та ін.).

В природних умовах людина розглядає предмети, які знаходяться поблизу неї і на різних відстанях включно до горизонту (розслабляючи при цьому м'язи ока). Крім того, має місце вільний рух очей у всі боки. Відтак функціонує все поле сітківки ока. Різноманітні м'язи ока і різноманітні ділянки поля сітківки функціонують поперемінно, отримуючи можливість відновлювати свій функціональний потенціал. Умови зорової роботи при використанні ВДТ набагато жорсткіші, оскільки у користувача комп'ютера "працює" лише ближній зір, тому елементи ока, що його забезпечують знаходяться у постійному напруженні.

2. Змінами умов, характерних для традиційного зорового процесу читання (темні знаки на світлому фоні при падаючому світловому потоці), а також демонстрування зображення на майже вертикальній поверхні, що випромінює світловий потік, а отже, потребує пониженого загального освітлення на робочому місці. В деяких випадках ВДТ відтворює яскраві знаки на темному фоні (зворотнє зображення затруднює адаптацію);

3. Світлотехнічною різноманітністю об'єктів зорової роботи що пов'язана з наявністю трьох об'єктів (екран, клавіатура, документація), розташованих у різних зонах спостереження, що вимагає багаторазового переведення лінії зору від одного до іншого.

Умови роботи з ВДТ ускладнюються необхідністю постійної перебудови апаратів акомодатії та конвергенції, не кажучи вже про постійну необхідність переадаптації від яскравих об'єктів з позитивним контрастом на темні — з негативним. Разом узяті всі ці особливості створюють багато незручностей, а також напруження м'язового та світловідчувачого апарату очей;

4. Робота з пульсуючим самосвітним об'єктом, який постійно перебуває у центрі поля зору, що не відповідає нормативним вимогам щодо обмеження пульсації та засліпленості. Наявність пульсації яскравості знаків викликає дискомфорт і втому, загальну й здорову;

5. Несприятливим розподілом яскравості у полі зору (стеля, стіни, меблі тощо можуть виявитися світлішими, ніж центр поля зору - темний, обмежено освітлений та іноді малозаповнений знаками екран монітора);

6. Засліплююча дія світильників, які освітлюють приміщення на робочому місці з ВДТ більша, ніж на інших, бо лінія зору користувача при роботі з екраном майже горизонтальна, що призводить до зменшення кута дії різних засліплюючих джерел (світильники, вікна і т. п.) і, відповідно, до зростання засліпленості.

Висновок.

Отже, порушення зорових функцій користувачів ВДТ пов'язані, головним чином, з чотирма групами факторів:

- параметрами освітлення робочого місця;
- характеристиками дисплея;
- специфікою роботи на ВДТ;
- неправильною організацією робочого місця.

Уникнення негативних факторів впливу на здоров'я і життя користувачів можливе лише при дотриманні вимог і норм охорони праці та правильної організації робочого розпорядку та ергономіки робочого місця.

ВИСНОВКИ

Основні наукові та практичні результати полягають в наступному.

1. У результаті проведеного аналізу наукових досліджень і статистичних даних щодо захворюваності на цукровий діабет як в Україні, так і в світі, встановлено що дане захворювання є досить розповсюдженим і зберігає тенденції до становлення як глобальної пандемії.

2. Виявлено основні фактори впливу на розвиток і появу підвищеного рівня глюкози в крові, що в основному залежить від специфіки життєдіяльності людини, віку та пізнього звернення у лікувальні установи, що дало змогу обґрунтувати доцільність та актуальність розробки і впровадження комп'ютерних систем виявлення, аналізу та прогнозування розвитку цукрового діабету.

3. Проаналізовано існуючі підходи, методи та інструменти для визначення рівня цукру в крові людини і встановлено, що найбільш поширеними для побутового використання є пристрої інвазивного та неінвазивного вимірювання, а системи моніторингу в Україні відсутні або недосконалі, що не дає змоги забезпечити якість обслуговування пацієнтів та виконувати прогнозування розвитку цього захворювання.

4. Розроблено алгоритми функціонування глюкометра та глобальної інформаційної системи управління медичними даними, які в комплексі становлять систему цілодобового моніторингу та керування рівнем глюкози в крові пацієнта, що дає змогу забезпечувати збір, опрацювання та прогнозування появи чи розвитку цукрового діабету.

5. Обґрунтовано можливі рішення щодо забезпечення розподіленості системи моніторингу та опрацювання даних захворюваністю цукровим діабетом, що дало змогу встановити оптимальний шлях побудови таких систем із застосуванням підходу змішаної фрагментації на вузлах мережі розподілу та управління навантаженням на основі програмного балансувальника.

6. Побудовано та математично представлено концептуальну модель розподіленої архітектури системи збору та опрацювання даних для моніторингу рівня цукру в крові, що включає в себе сукупність локальних і центрального вузла управління і дає змогу обмінюватися повідомленнями та виконувати прогнозування розвитку хвороби. Основна функція локальних вузлів у розподіленій системі полягає у зборі даних з глюкометрів та передачі їх у центральний вузол управління. Найбільш важлива функція центрального вузла управління полягає в агрегації даних від локальних вузлів за допомогою дистриб'ютора транзакцій, сервісної шини та агрегатора, що дозволяє забезпечувати цілісність даних та подальше прогнозування рівня цукру в крові.

7. Проведено аналіз даних, які включають 24 фактори впливу на появу і розвиток цукрового діабету, обґрунтовано використання мови програмування Python та відповідних відкритих бібліотек, які в подальшому забезпечать моделювання процесу опрацювання даних, визначення факторів з найбільшою вагою впливу на розвиток цукрового діабету та прогнозування схильності людей до цього захворювання .

8. Реалізовано процедури визначення кореляції між цільовою змінною та іншими незалежними змінними, наявними у даних, що дало змогу встановити, що найбільше появу і розвиток різного типу діабету провокує куріння, алкоголь, психологічний стан і доходи, які одержує людина, а в комплексі факторів – паління та надмірне споживання алкоголю, інсульт та серцево-судинні хвороби чи інфаркти, високий артеріальний тиск і рівень холестерину.

9. Реалізовано алгоритми прогнозування (класифікації) розвитку цукрового діабету на основі підходів дерев прийняття рішень, випадкового лісу, логістичної регресії та XGBoost і встановлено, що при врахуванні усіх факторів впливу на розвиток цукрового діабету найбільшу точність прогнозування забезпечує XGBoost на рівні 83%. Даний алгоритм застосовано після вибору 14 найбільш впливових факторів і при цьому точність підвищилась до інтервалу 98-100%.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Статистика захворювання цукровим діабетом. URL: <https://www.who.int/diabetes/> (дата звернення 10.11.2022 р.).
2. Карамишева Т. В. Діабет. Сучасна енциклопедія з рекомендаціями / Тетяна Вікторівна Карамишева. Київ: Ексмо, 2016. 464 с.
3. Ахманов М. Настольна книга діабетика / М. Ахманов, Х. Астамірова. – Київ: Ексмо, 2015. 496 с.
4. Діагностика та контроль цукрового діабету. URL: <https://indar.com.ua/ua/node/46> (дата звернення 11.11.2022 р.).
5. Неінвазивні глюкометри. URL: <http://ukrdiagnostika.ua/press/portable-glucometer.html> (дата звернення 13.11.2022 р.)
6. Segman Y. J. Device and Method for Noninvasive Glucose Assessment. URL: <https://pubmed.ncbi.nlm.nih.gov/29575926/>.
7. Python Tutorial. URL: <https://www.w3schools.com/python/default.asp> (дата звернення 15.11.2022 р.).
8. Pandas documentation. URL: <https://pandas.pydata.org/docs/index.html> (дата звернення 28.11.2022 р.).
9. Y. Bengio, R. Ducharme, and P. Vincent. A neural probabilistic language model. In *Advances in Neural Information Processing Systems 13 (NIPS 2000)*. 2001. p. 932–938.
10. Graves and J. Schmidhuber. Framewise phoneme classification with bidirectional LSTM networks. In *2005 International Joint Conference on Neural Networks (ICJNN'05)*. 2005. p. 23–43.
11. Preprocessing data. URL: <https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing> (дата звернення 05.11.2022 р.).
12. API reference. URL: <https://pandas.pydata.org/docs/reference/index.html> (дата звернення 10.11.2022 р.).
13. Python-recsys on Github. URL: <https://github.com/ocelma/python-recsys>

14. Preprocessing data. URL: <https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing> (дата звернення 10.11.2022 р.)

15. G. Guido, A. Vitale, V. Astarita, F. Saccomanno, V. P. Giofr e, and V. Gallelli, “Estimation of safety performance measures from smartphone sensors,” *Procedia—Social and Behavioral Sciences*, vol. 54, pp. 2012. 1095–1103.

16. Луцкiв А.М., Баран С.О. Технологiї неiнвaзивного вимiрювання рiвня глюкози в кровi. Матерiали Х науково-технiчної конференцiї Тернопiльського нацiонального технiчного унiверситету iменi Iвана Пулюя «Iнформацiйнi моделi, системи та технологiї» (8-9 грудня 2022 року). Тернопiль: ТНТУ. 2022. С. 114.

17. Луцкiв А.М., Баран С.О. Алгоритми машинного навчання для прогнозування рiвня глюкози в кровi. Матерiали Х науково-технiчної конференцiї Тернопiльського нацiонального технiчного унiверситету iменi Iвана Пулюя «Iнформацiйнi моделi, системи та технологiї» (8-9 грудня 2022 року). Тернопiль: ТНТУ. 2022. С. 114.

18. W. J. Zhang, S. X. Yu, Y. F. Peng, Z. J. Cheng, and C. Wang, “Driving habits analysis on vehicle data using error backpropagation neural network algorithm,” in *Computing, Control, Information and Education Engineering*, vol. 55, CRC Press, Guilin, China, 2015.

19. H. Drucker, J. C. Chris, L. Kaufman, A. Smola, and V. Vapnik, “Support vector regression machines,” in *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, USA. pp. 155–161, 1997.

20. H.-l. Feng, “Study on prediction model of ecological security index in Chongqing city based on SVR model,” *Computer Science*, vol. 40, no. 8. 2013. pp. 245–248.

21. Z. Ramedani, M. Omid, A. Keyhani, S. Shamshirband, and B. Khoshnevisan, “Potential of radial basis function based support vector regression for global solar radiation prediction,” *Renewable and Sustainable Energy Reviews*, vol. 39. 2014. pp. 1005–1011

22. API reference. URL: <https://pandas.pydata.org/docs/reference/index.html> (дата звернення 10.11.2022 р.).

23. NumPy Reference. URL: <https://numpy.org/doc/stable/reference/index>.
24. Garbade M. J. Understanding K-means Clustering in Machine Learning. URL: <https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1>
25. Катренко Л.А., Катренко А.В. Охорона праці в галузі комп'ютерингу. Львів: Магнолія-2006. 2012. 544 с.
26. Желібо Є. Безпека життєдіяльності. К.: 2001. 483 с.

Додаток А
Тези конференцій

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ**

МАТЕРІАЛИ

X НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

**«ІНФОРМАЦІЙНІ МОДЕЛІ,
СИСТЕМИ ТА ТЕХНОЛОГІЇ»**



7–8 грудня 2022 року

**ТЕРНОПІЛЬ
2022**

А. Станько АНАЛІЗ КОНЦЕПЦІЇ ВСЕОСЯЖНОГО ІНТЕРНЕТУ – ІоЕ	
A. Stanko ANALYSIS OF THE CONCEPT OF THE INTERNET OF EVERYTHING – ІоЕ	53
М. Турчиняк ТЕХНОЛОГІЇ ВПЛИВУ СОЦІАЛЬНИХ МЕРЕЖ НА ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ	
M. Turchyniak TECHNOLOGIES OF THE INFLUENCE OF SOCIAL NETWORKS ON ENSURING INFORMATION SECURITY	55
Д. Урбан АНАЛІЗ ЗАГРОЗ КОМП'ЮТЕРНИХ СИСТЕМ	
D. Urban ANALYSIS OF COMPUTER SYSTEM THREATS	57
А. Хом'як СИГНАЛИ ГОЛОВНОГО МОЗКУ, ЯКІ МОЖНА ОТРИМАТИ НЕІНВАЗИВНИМИ МЕТОДАМИ	
A. Khomiak BRAIN SIGNALS OBTAINABLE VIA NON-INVASIVE IMAGING	58
Г. Шимчук, О. Голотенко, Р. Золотий ОСНОВНІ ПРОБЛЕМИ ТА ЗАГРОЗИ ХМАРНОЇ БЕЗПЕКИ	
G. Shymchuk, O. Holotenko, R. Zoloty USE THE MAIN PROBLEMS AND THREATS OF CLOUD SECURITY	59
А. Мачужак АВТОМАТИЗАЦІЯ ЗАДАЧ ТЕСТУВАННЯ ТА РОЗГОРТАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ	
A. Machuzhak AUTOMATION OF SOFTWARE TESTING AND DEPLOYMENT TASKS	61
СЕКЦІЯ 3. КОМП'ЮТЕРНІ СИСТЕМИ ТА МЕРЕЖІ	
М. Домарецький ОГЛЯД СИСТЕМ ДЛЯ РОЗПІЗНАВАННЯ ЖЕСТІВ	
M. Domaretskyi REVIEW OF GESTURE RECOGNITION SYSTEMS	62
А. Луцків, С. Баран ТЕХНОЛОГІЇ НЕІНВАЗИВНОГО ВИМІРЮВАННЯ РІВНЯ ГЛЮКОЗИ В КРОВІ	
A. Lutskiv, S. Baran TECHNOLOGIES OF NON-INVASIVE GLUCOSE LEVEL MEASUREMENT IN BLOOD	63
А. Луцків, С. Баран АЛГОРИТМИ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ РІВНЯ ГЛЮКОЗИ В КРОВІ	
A. Lutskiv, S. Baran MACHINE LEARNING ALGORITHMS FOR PREDICTING THE LEVEL OF GLUCOSE IN THE BLOOD	64
А. Луцків, М. Бондаренко ОСОБЛИВОСТІ ЗАДАЧ І ФУНКЦІЙ DEVOPS ФАХІВЦІВ	
A. Lutskiv, M. Bondarenko FEATURES OF TASKS AND FUNCTIONS OF DEVOPS SPECIALISTS	65

УДК 004.3

А. Луцків, С. Баран

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ТЕХНОЛОГІЇ НЕІНВАЗИВНОГО ВИМІРЮВАННЯ РІВНЯ ГЛЮКОЗИ В КРОВІ

UDC 004.3

A. Lutskiv, S. Baran

TECHNOLOGIES OF NON-INVASIVE GLUCOSE LEVEL MEASUREMENT IN BLOOD

Суть функціонування неінвазивних вимірювачів глюкози полягає в аналізі за станом кровоносних судин, тобто передбачаються не прямі вимірювання. Окрім цього, існує декілька функціональних моделей, які забезпечують обчислення концентрації глюкози у крові на основі аналізу стану шкірного покриву. Для цього достатньо забезпечити контакт пристрою з ділянкою частини тіла людини. Загалом за технологіями неінвазивного визначення рівня цукру в крові використовуються технології, які показано на рис. 1.

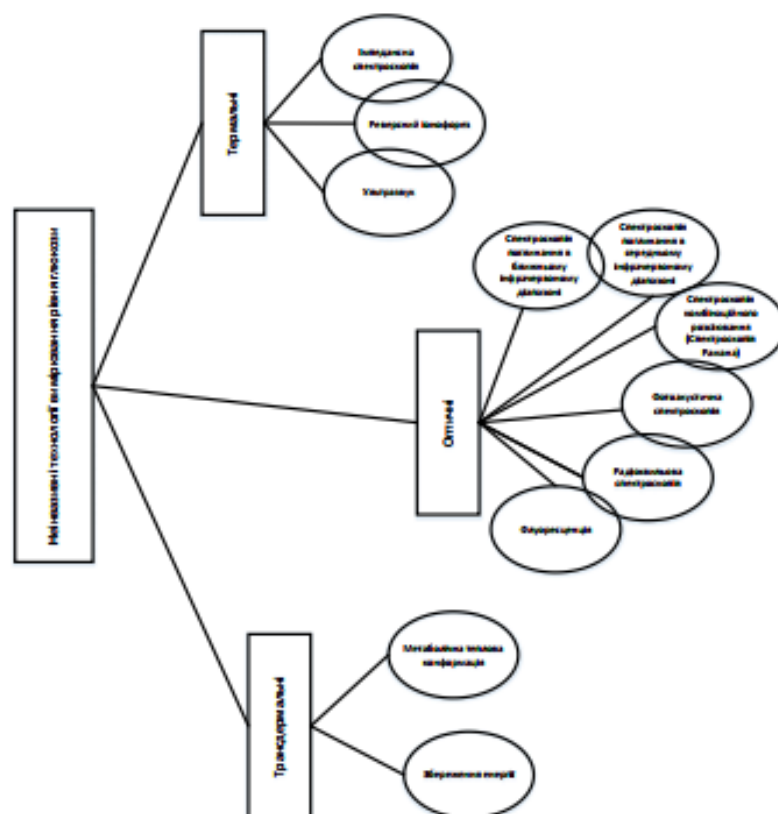


Рисунок 1. Технології неінвазивного визначення рівня глюкози в крові

Оптичні методи використовують різні властивості світла для взаємодії з глюкозою в залежності від її концентрації. Трансдермальний метод передбачає вимірювання рівня глюкози через шкіру за допомогою електричних імпульсів або ультразвуку.

УДК 004.8

А. Луцків, С. Баран

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

АЛГОРИТМИ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ РІВНЯ ГЛЮКОЗИ В КРОВІ

UDC 004.8

A. Lutskiv, S. Baran

MACHINE LEARNING ALGORITHMS FOR PREDICTING THE LEVEL OF GLUCOSE IN THE BLOOD

Для вибору та проведення експериментів щодо прогнозування рівня глюкози в крові для трьох категорій людей: ті, у яких нормальний рівень цукру в крові, ті, хто хворіє діабетом першого типу, ті, хто хворіє діабетом другого типу, потрібно застосувати методи машинного навчання, оскільки кількість даних є надзвичайно велика, а визначити і врахувати кореляції між факторами, або їх групами вручну практично неможливо.

Потрібно відмітити, що при прогнозуванні рівня цукру в крові необхідно розв'язати задачу класифікації та визначити сукупність і пріоритет факторів, які найбільше впливають на розвитку цукрового діабету. У даному випадку, прогнозування трактується як prediction, а не прогнозування розвитку цукрового діабету в часі – forecast. Як приклад задач, методів і алгоритмів, для яких можна застосовувати інструменти машинного навчання належать такі, які проілюстровані на рис. 1.

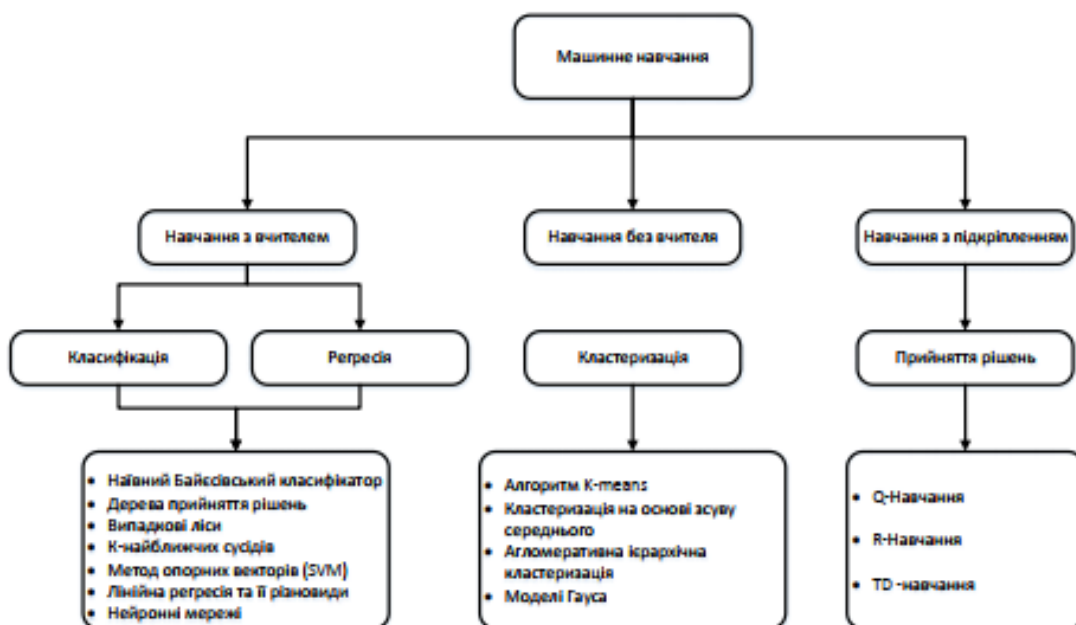


Рисунок 1. Застосування методів і алгоритмів машинного навчання

З практичної точки зору, для розв'язку задачі такого класу найбільш ефективними є алгоритми: логістичної регресії; дерев прийняття рішення; XGBoost; «випадкового лісу». При проведенні експериментів найкращі результати показав алгоритм XGBoost, що забезпечив точність на рівні більше 90%.