

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова
праця на правах рукопису

БОЛОБАН ОЛЕГ АНАТОЛІЙОВИЧ

УДК 004.046; 004.896; 004.021

ДИСЕРТАЦІЯ

**МЕТОДИ ТА ЗАСОБИ ПРОГНОЗУВАННЯ ЗАХВОРЮВАНЬ ДИХАННЯ
ТА АПНОЕ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ ТА МІКРО
СЕРВІСНОЇ АРХІТЕКТУРИ**

122 – Комп'ютерні науки

12 – Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії.

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

_____ О.А. Болобан

Науковий керівник:
ПЕТРЕНКО Анатолій Іванович
доктор технічних наук, професор

КИЇВ – 2025

АНОТАЦІЯ

Болобан О.А. МЕТОДИ ТА ЗАСОБИ ПРОГНОЗУВАННЯ ЗАХВОРЮВАНЬ ДИХАННЯ ТА АПНОЕ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ ТА МІКРО СЕРВІСНОЇ АРХІТЕКТУРИ. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 122 Комп'ютерні науки за спеціальністю 122 Комп'ютерні науки. – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, 2025.

Синдром обструктивного апное сну та хропіння є поширеними патологіями, що значно впливають на якість життя пацієнтів, особливо після пандемії COVID-19. Розлади сну можуть спричиняти сонливість, зниження концентрації, когнітивні порушення та підвищувати ризик серцево-судинних захворювань і смертності. Однак лише 20% пацієнтів отримують вчасний діагноз, оскільки існуючі засоби діагностики є дорогими, малодоступними або не забезпечують комплексного аналізу дихальних параметрів. Більшість комерційних пристроїв обмежені лише вимірюванням рівня кисню в крові та серцевого ритму, що не дозволяє ефективно ідентифікувати дихальні патології.

Дане дослідження спрямоване на розробку інтелектуальної системи моніторингу та прогнозування порушень дихання під час сну на основі штучного інтелекту. Запропонована система поєднує технології машинного навчання, нейронних мереж та методів обробки біомедичних сигналів для виявлення апное та інших дихальних розладів. Архітектура рішення базується на мікросервісному підході з використанням розподілених обчислень, а саме прикінцевих обчислень, що забезпечує ефективну обробку даних безпосередньо на пристрої та мінімізує затримки передачі.

Запропонований підхід передбачає адаптивну обробку сигналів фотоплетизмограми (англ. Photoplethysmography (PPG)) та інших біомедичних показників з використанням методів фільтрації шумів, а також прогнозування

стану пацієнта за допомогою моделі довга короткочасна пам'ять (англ. Long short-term memory (LSTM)). Інтеграція сервісно-орієнтованої архітектури дозволяє масштабувати систему для подальшого розвитку, що робить її придатною для персонального та клінічного використання.

Результати дослідження відповідають сучасним тенденціям розвитку цифрової медицини, сприяють впровадженню інноваційних технологій у сфері діагностики розладів сну та можуть бути використані в програмах розвитку «Інтернету медичних речей» (англ. Internet of Medical Things (IoMT)) та когнітивних обчислень. Метою дисертації є розробка моделі класифікатора апное-гіпноное сну та впровадження в мікросервісну архітектуру, забезпечуючи ефективну класифікацію типу апное-гіпноное сну в реальному часі.

Об'єкт дослідження – процеси збору, обробки та прогнозування респіраторних захворювань на основі біомедичних сигналів із використанням штучного інтелекту та мікро сервісної архітектури.

Предмет дослідження – методи фільтрації, аналізу та прогнозування респіраторних захворювань на основі сигналів фотоплетизмографії, нейронних мережових алгоритмів та технологій мікросервісної архітектури для забезпечення ефективного моніторингу та автоматизованої діагностики.

Наукова новизна отриманих результатів полягає у тому, що в дисертаційній роботі:

1. Вперше запропоновано модель прогнозування респіраторних патологій, яка використовує сигнали фотоплетизмографії та методи глибокого навчання для оцінки стану дихальної системи. Відмінністю запропонованого підходу є застосування адаптивної моделі довгої короткочасної пам'яті, що аналізують часові ряди біомедичних сигналів та виявляють закономірності, пов'язані з апное та іншими порушеннями дихання. На відміну від традиційних методів, які базуються на простому порівнянні порогових значень кисню в крові, запропонована модель дозволяє враховувати багатофакторний аналіз змін сигналу та його динаміку, що суттєво підвищує точність виявлення патологій.

2. Модифіковано підхід до фільтрації медичних даних та реалізовано комбінований метод обробки фотоплетизмографічних сигналів. Вперше запропоновано вдосконалений підхід до фільтрації біомедичних сигналів, що включає комбіноване використання адаптивного фільтра Калмана, середньозваженого фільтра та алгоритмів машинного навчання для очищення сигналів у реальному часі. Особливістю підходу є динамічне налаштування параметрів фільтрації залежно від характеристик шуму та артефактів, що виникають під час рухової активності або зміни умов вимірювання. Проведений аналіз показав, що запропонований метод перевершує традиційні фільтри нижніх частот, верхніх частот та ковзного середнього за точністю відновлення сигналу та стабільністю його обробки у змінних умовах.

Практичне значення отриманих результатів полягає у наступному:

1. Розроблена інтегрована система для відстеження та прогнозування дихальних патологій за межами медичних закладів. Запропонована система забезпечує можливість моніторингу стану дихальної системи пацієнтів у домашніх умовах, що значно покращує доступність діагностики. Використання мікросервісної архітектури дозволяє легко інтегрувати нові сенсори та моделі прогнозування, що робить систему гнучкою та масштабованою для різних груп користувачів.

2. Розроблена система моніторингу і класифікації стадій апное під час сну на основі штучного інтелекту для домашнього використання, яка дозволяє попереджати пацієнтів про ризики захворювання апное під час сну самостійно. Запропоноване рішення надає користувачам можливість безперервного моніторингу показників дихання та їх аналізу за допомогою штучного інтелекту. Система не лише відображає поточні параметри, але й проводить аналіз тенденцій та класифікує можливі стадії апное сну, що дає змогу пацієнтам своєчасно звертатися до лікаря.

3. Система підтримує прийняття діагностичних рішень шляхом автоматичної класифікації подій апное за 30-секундними вікнами сигналів фотоплетизмографії, сатурації та ЧСС, надаючи лікарю результати класифікації для аналізу, корекції та

подальшого визначення необхідності стаціонарного догляду за пацієнтом. Завдяки інтеграції системи з медичними інформаційними платформами лікарі отримують можливість доступу до історії змін стану пацієнтів, що допомагає у прийнятті обґрунтованих рішень щодо діагностики та лікування. Віддалений моніторинг дозволяє зменшити навантаження на клініки та покращити якість обслуговування пацієнтів.

Результати проведеного дослідження були представлені у провідних наукових виданнях, що підтверджує їхню актуальність і наукову новизну:

- Наукова стаття Boloban, O., Pysmennyi, I., Kyslyi, R., & Kyriusha, B. (2024). Development of a patient health monitoring system based on a service-oriented architecture using artificial intelligence. *Technology Audit and Production Reserves*, 3(2(77)), P. 23–29.
- Наукова стаття Петренко, А. І., Болобан, О. А. (2025). Розробка інфраструктури для інтеграції штучного інтелекту в телемедичні системи за допомогою мікросервісної архітектури. *Таврійський науковий вісник. Серія: Технічні науки*, (1), С. 148-158.
- Наукова стаття Болобан О.А., Петренко А.І. Особливості застосування алгоритму EMD для оцінки рівня кисню в крові. *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки. Том 34 (73). № 2. 2023. С. 125-130.*
- Наукова стаття Петренко, А., Болобан, О. (2023). Узагальнення інформації з прикладами про можливості використання сервіс-орієнтованого підходу та технологій штучного інтелекту в галузі e-health. *Технологічний аудит та резерви виробництва*, 4 (2 (72)), С. 10-17.
- Наукова стаття Петренко, А., Болобан, О. (2024). Дослідження ефективності алгоритмів фільтрації біомедичних сигналів у реальному часі. *Інфокомунікаційні та комп'ютерні технології*, 2(08), С. 53-61.

Ключові слова: машинне навчання, нейронні мережі, класифікація, інтернет речей, енергоефективні обчислення, формальна специфікація, розподілені система, апное та гіпопное сну, мікросервіси, COVID-19, моделювання, скорочення часу виконання, часова логіка дій.

ANOTATION

Boloban O.A. METHODS AND TOOLS FOR PREDICTING RESPIRATORY DISEASES AND APNEA BASED ON ARTIFICIAL INTELLIGENCE AND MICRO SERVICE ARCHITECTURE – Qualification Scientific Work in the Form of a Manuscript.

Dissertation for the Degree of Doctor of Philosophy in the field of knowledge 122 Computer Science, specialty 122 Computer Science. – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, 2025.

Obstructive sleep apnea syndrome and snoring are common pathologies that significantly impact patients' quality of life, especially after the COVID-19 pandemic. Sleep disorders can cause drowsiness, decreased concentration, cognitive impairments, and an increased risk of cardiovascular diseases and mortality. However, only 20% of patients receive a timely diagnosis, as existing diagnostic tools are expensive, inaccessible, or do not provide a comprehensive analysis of respiratory parameters. Most commercial devices are limited to measuring blood oxygen levels and heart rate, which do not effectively identify respiratory pathologies.

This study aims to develop an intelligent monitoring and prediction system for respiratory disorders during sleep based on artificial intelligence. The proposed system integrates machine learning technologies, neural networks, and biomedical signal processing methods to detect apnea and other respiratory disorders. The solution architecture is based on a microservice approach using edge computing, ensuring efficient data processing directly on the device and minimizing transmission delays.

The proposed approach involves adaptive processing of photoplethysmogram (PPG) signals and other biomedical indicators using noise filtering methods, as well as predicting patient conditions with Long short-term memory (LSTM) models. The integration of a service-oriented architecture allows for system scalability and further development, making it suitable for both personal and clinical use.

The research results align with modern trends in digital medicine, fostering the implementation of innovative technologies in sleep disorder diagnostics. They can be applied in the development programs of the Internet of Medical Things (IoMT) and cognitive computing. The aim of the dissertation is to develop a sleep apnea–hypopnea classifier model and integrate it into a microservices architecture, ensuring efficient real-time classification of sleep apnea–hypopnea types.

The object of the study is the processes of collecting, processing, and predicting respiratory diseases based on biomedical signals using artificial intelligence and a microservice architecture.

The subject of the study is the methods of filtering, analyzing, and predicting respiratory diseases based on PPG signals, neural network algorithms, and microservice architecture technologies to ensure efficient monitoring and automated diagnostics.

The scientific novelty of the obtained results lies in the fact that in the dissertation:

1. For the first time, a model for predicting respiratory pathologies has been proposed, utilizing PPG signals and deep learning methods to assess the condition of the respiratory system. The distinguishing feature of the proposed approach is the use of adaptive LSTM models that analyze time-series biomedical signals and identify patterns associated with apnea and other breathing disorders. Unlike traditional methods, which rely on simple threshold comparisons of blood oxygen levels, the proposed model incorporates a multifactorial analysis of signal variations and their dynamics, significantly improving the accuracy of pathology detection.

2. The approach to filtering medical data has been modified, and a combined method for processing PPG signals has been implemented. For the first time, an advanced approach to biomedical signal filtering has been proposed, incorporating the combined use of an adaptive Kalman filter, a weighted moving average filter, and machine learning algorithms for real-time signal purification. A key feature of this approach is the dynamic adjustment of filtering parameters based on noise characteristics and artifacts that arise due to motion activity or changes in measurement conditions. The conducted analysis demonstrated that the proposed method outperforms traditional low pass, high pass,

moving average and Kalman filters in terms of signal recovery accuracy and processing stability under variable conditions.

The practical significance of the obtained results is as follows:

1. A fully integrated system for tracking and predicting respiratory pathologies outside medical facilities has been developed. The proposed system enables the monitoring of patients' respiratory health in home environments, significantly improving diagnostic accessibility. The use of a microservice architecture allows for easy integration of new sensors and predictive models, making the system flexible and scalable for different user groups.

2. A home-use monitoring and prediction system for respiratory disorders during sleep based on artificial intelligence has been developed, allowing patients to independently monitor their health. The proposed solution enables continuous monitoring of respiratory parameters and their analysis using artificial intelligence. The system not only displays real-time data but also analyzes trends and predicts potential risks, helping patients seek medical attention in a timely manner.

3. The developed monitoring and prediction system for respiratory disorders during sleep based on artificial intelligence supports doctors in diagnosing respiratory disorders through remote access to patient information. By integrating with medical information platforms, physicians can access the history of patient condition changes, aiding in informed decision-making for diagnosis and treatment. Remote monitoring helps reduce the burden on clinics and improves the quality of patient care.

The results of the conducted research have been presented and discussed in leading scientific publications, confirming their relevance and scientific novelty:

- Scientific article Boloban, O., Pysmennyi, I., Kyslyi, R., & Kyriusha, B. (2024). Development of a patient health monitoring system based on a service-oriented architecture using artificial intelligence. *Technology Audit and Production Reserves*, 3(2(77)), P. 23-29.
- Scientific article Petrenko, A. I., Boloban, O. A. (2025). Development of infrastructure for the integration of artificial intelligence into telemedicine systems

using microservice architecture. Tavian Scientific Bulletin. Series: Technical Sciences, (1), P. 148-158.

- Scientific article Boloban O.A., Petrenko A.I. Features of the EMD algorithm application for estimation of blood oxygen level. Scientific notes of Vernadsky TNU. Series: Technical sciences. Vol. 34 (73). № 2. 2023. P. 125-130.
- Scientific article Petrenko, A., Boloban, O. (2023). Generalization of information with examples on the possibilities of using a service-oriented approach and artificial intelligence technologies in the field of e-health. Technological audit and production reserves, 4 (2 (72)), P. 10-17.
- Scientific article Petrenko, A., Boloban, O. (2024). Investigation of the effectiveness of real-time filtering algorithms for biomedical signals. Infocommunication and computer technologies, 2(08), P. 53-61.

Keywords: machine learning, neural networks, classification, Internet of Things (IoT), energy-efficient computing, formal specification, distributed systems, sleep apnea and hypopnea, microservices, COVID-19, modeling, execution time reduction, Temporal Logic of Actions (TLA).

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

1. Boloban, O., Pysmennyi, I., Kyslyi, R., & Kyriusha, B. (2024). Development of a patient health monitoring system based on a service-oriented architecture using artificial intelligence. *Technology Audit and Production Reserves*, 3(2(77)), P. 23–29. (Фахове видання категорії “Б”) (Здобувачу належить реалізація програмної частини, порівняльний аналіз запропонованих методів)
2. Петренко, А. І., Болобан, О. А. (2025). Розробка інфраструктури для інтеграції штучного інтелекту в телемедичні системи за допомогою мікросервісної архітектури. *Таврійський науковий вісник. Серія: Технічні науки*, (1), С. 148-158. (Фахове видання категорії “Б”) (Здобувачу належить аналіз літературних джерел, аналіз та порівняння існуючих підходів та методів, реалізація програмної частини, обґрунтування та узагальнення отриманих результатів)
3. Болобан О.А., Петренко А.І. Особливості застосування алгоритму EMD для оцінки рівня кисню в крові. *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки. Том 34 (73). № 2. 2023. С. 125-130.* (Фахове видання категорії “Б”) (Здобувачу належить постановка задачі, аналіз літературних джерел, аналіз та порівняння існуючих підходів та методів)
4. Петренко, А., Болобан, О. (2023). Узагальнення інформації з прикладами про можливості використання сервіс-орієнтованого підходу та технологій штучного інтелекту в галузі e-health. *Технологічний аудит та резерви виробництва*, 4 (2 (72)), С. 10-17. (Фахове видання категорії “Б”) (Здобувачу належить постановка задачі, аналіз наявних рішень, реалізація програмної частини, обґрунтування та узагальнення отриманих результатів)
5. Петренко, А., Болобан, О. (2024). Дослідження ефективності алгоритмів фільтрації біомедичних сигналів у реальному часі. *Інфокомунікаційні та комп’ютерні технології*, 2(08), С. 53-61. (Фахове видання категорії “Б”) (Здобувачу належить аналіз літературних джерел, аналіз та порівняння існуючих підходів та методів, реалізація програмної частини, обґрунтування та узагальнення отриманих результатів)

Перелік умовних позначень та скорочень

| | |
|--------------|--|
| AI | Artificial Intelligence (штучний інтелект) |
| API | Application Programming Interface |
| AHI | Apnea–Hypopnea Index (індекс апное-гіпноное) |
| BLE | Bluetooth Low Energy (бездротовий інтерфейс низького енергоспоживання) |
| CNN / 1D-CNN | Convolutional Neural Network / One-Dimensional Convolutional Neural Network |
| COVID-19 | Coronavirus Disease 2019 |
| EMD | Empirical Mode Decomposition (емпіричне модове розкладання) |
| ESP32-C3 | модель мікроконтролера ESP32 серії C3 |
| GATT | Generic Attribute Profile (протокол для BLE-сервісів) |
| GDPR | General Data Protection Regulation (Загальний регламент захисту даних) |
| HIPAA | Health Insurance Portability and Accountability Act (Закон про переносимість і підзвітність медичного страхування) |
| HR (ЧСС) | Heart Rate (частота серцевих скорочень) |
| IoMT | Internet of Medical Things (Інтернет медичних речей) |
| JSON | JavaScript Object Notation |
| JSONB | Binary JSON (тип даних у PostgreSQL для зберігання JSON-структур) |
| JWT | JSON Web Token (механізм авторизації через токени) |
| LSTM | Long Short-Term Memory (рекурентна нейронна мережа з довгою короткочасною пам'яттю) |
| MQTT | Message Queuing Telemetry Transport (легкий протокол обміну повідомленнями, зазвичай для IoT) |
| PPG (ФПГ) | Photoplethysmography (фотоплетизмографія) |
| REST API | Representational State Transfer API (стиль побудови веб-інтерфейсів) |
| SOA | Service-Oriented Architecture (сервіс-орієнтована архітектура) |

SQL Structured Query Language система управління базами даних
PostgreSQL
TLA Temporal Logic of Actions (мова часової логіки дій для формальної
специфікації)
TLC TLA Checker (інструмент перевірки моделей у TLA+)

ЗМІСТ

| | |
|--|----|
| ВСТУП..... | 17 |
| РОЗДІЛ 1. ВПРОВАДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ В ОХОРОНУ ЗДОРОВ'Я: МОЖЛИВОСТІ ТА ВИКЛИКИ | 30 |
| РОЗДІЛ 2. МЕТОДОЛОГІЧНИЙ АНАЛІЗ ПРОБЛЕМИ ДІАГНОСТИКИ АПНОЕ СНУ З ВИКОРИСТАННЯМ ГЛИБОКОГО НАВЧАННЯ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ДІАГНОСТИЧНИХ РІШЕНЬ ЛІКАРЕМ | 33 |
| 2.1 Захворювання органів дихання та причини порушення дихання та виникнення апное сну..... | 33 |
| 2.2 Традиційні клінічні методи реєстрації та аналізу біомедичних сигналів для діагностики апное сну..... | 34 |
| 2.3 Застосування методів цифрової обробки сигналів для покращення якості діагностики апное сну..... | 36 |
| 2.4 Аналіз застосування методів глибокого навчання до задачі класифікації сигналів фотоплетизмограм для виявлення апное сну | 37 |
| 2.4.1. Виділення репрезентативних ознак, які можуть бути пов'язані з апное сну | 38 |
| 2.4.2. Визначення джерел даних..... | 42 |
| 2.4.3. Попередня обробка сигналів фотоплетизмограми. | 47 |
| РОЗДІЛ 3. РОЗРОБКА МІКРО СЕРВІСНОЇ АРХІТЕКТУРИ СИСТЕМИ МОНІТОРИНГУ І ПРОГНОЗУВАННЯ ДИХАЛЬНИХ РОЗЛАДІВ ПІД ЧАС СНУ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ..... | 58 |
| 3.1. Мікросервісна архітектура..... | 60 |
| 3.2. Подієво-орієнтована архітектура | 62 |
| 3.3. Прикінцеві обчислення | 65 |
| 3.4 Комунікація між компонентами | 69 |
| 3.4.1 Bluetooth низького енергоспоживання..... | 69 |
| 3.4.2. Асинхронна передача повідомлень між мікросервісами. | 70 |
| 3.4.3. Використання інтерфейсу прикладного програмування для взаємодії між мобільним додатком та ІІП шлюзу..... | 70 |
| 3.5. Розробка архітектури системи моніторингу і прогнозування дихальних розладів під час сну на основі штучного інтелекту | 74 |
| 3.5.1. Збір даних | 75 |
| 3.5.2 Обробка даних на рівні периферії..... | 75 |
| 3.5.3 Відправка даних в хмару..... | 75 |
| 3.5.4. Зберігання даних в хмарі | 76 |

| | |
|--|------------|
| | 15 |
| 3.5.5. Обробка даних в хмарі | 76 |
| 3.5.6. Підготовка даних для відправки до сервісу класифікації | 77 |
| 3.5.7. Обробка результатів | 77 |
| 3.5.8. Збереження результатів | 78 |
| 3.5.9. Відправка даних на мобільний пристрій | 78 |
| 3.6. Опис архітектури | 79 |
| РОЗДІЛ 4. ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ НАВЧАННЯ МОДЕЛЕЙ ДЛЯ КЛАСИФІКАЦІЇ АПНОЕ СНУ | 83 |
| 4.1 Вибір та реалізація алгоритмів машинного навчання для прогнозування захворювань дихання та апное | 84 |
| 4.1.1. Опис репозиторію цифрових сигналів фотоплетизмограм для розробки методів діагностики апное сну | 90 |
| 4.1.2. Підготовка даних для застосування методів глибокого навчання до задачі класифікації сигналів фотоплетизмограм для виявлення апное сну | 93 |
| 4.2. Порівняння результатів машинного навчання для різних архітектур нейронних мереж | 96 |
| 4.3. Навчання моделей на підготовлених даних, оптимізація гіпер параметрів, оцінка якості моделей за допомогою метрик | 107 |
| 4.4. Вибір найкращої моделі машинного навчання для діагностики апное сну .. | 109 |
| 4.5. Інтеграція моделей в мікросервісну архітектуру, забезпечення можливості їх оновлення та перенавчання | 110 |
| РОЗДІЛ 5. РОЗРОБКА НОСИМОГО ПРИСТРОЮ ДЛЯ РЕЄСТРАЦІЇ ТА ЗБОРУ ФПГ СИГНАЛУ | 113 |
| 5.1. Безпека передачі даних від носимого пристрою до мобільного | 116 |
| 5.2. Оптимізація програмного забезпечення для забезпечення енергоефективності та скорочення часу виконання | 118 |
| 5.3. Формальна специфікація та перевірка часової логіки системи засобами часової логіки дій | 120 |
| РОЗДІЛ 6. РЕАЛІЗАЦІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ДІАГНОСТИЧНИХ РІШЕНЬ ДЛЯ ВИЯВЛЕННЯ АПНОЕ СНУ | 122 |
| 6.1. Розробка енергоефективної прошивки ESP32-C3 для зчитування даних з ФПГ сенсору, фільтрація сигналів і шифрованої BLE-передачі сегментів. | 122 |
| 6.2. Створення мобільного застосунку для приймання BLE-даних, відображення телеметрії та ретрансляції сегментів у хмару | 126 |
| 6.3. Проектування мікро сервісного бекенду для приймання, зберігання та потокової обробки біомедичних даних | 140 |
| 6.4. Реалізація скрипту підготовки даних і навчання LSTM-класифікатора на сегментах HR та SpO ₂ | 143 |
| 6.5. Інтеграція натренованої LSTM-моделі в хмарну архітектуру за допомогою Amazon SageMaker Hosting та автоматичне масштабування AI-сервісу. | 146 |

| | |
|----------------------------------|-----|
| ВИСНОВОК | 148 |
| СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ | 150 |

ВСТУП

Актуальність теми.

Актуальність роботи зумовлена потребою у вирішенні проблеми автоматизації досліджень та вдосконалення інформаційно-алгоритмічного забезпечення систем моніторингу і прогнозування дихальних розладів під час сну на основі штучного інтелекту для виявлення апное сну та визначення стадії розвитку хвороби на основі аналізу сигналів фотоплетизмограм.

Актуальність дослідження обумовлена труднощами точного й безперешкодного збору даних для діагностики апное. Існуючі системи часто обмежують рухливість пацієнтів і вимагають використання громіздких пристроїв, що робить їх непридатними для щоденного застосування. Використання алгоритмів штучного інтелекту дозволяє оптимізувати процес діагностики, зменшуючи обсяг необхідних даних для обробки та забезпечуючи можливість прогнозування захворювань дихальної системи на ранніх стадіях. Мікросервісна архітектура системи сприяє її масштабованості, забезпечуючи інтеграцію нових функцій, таких як аналітичні модулі чи додаткові фільтри, без необхідності зміни основного програмного забезпечення. Це забезпечує гнучкість і адаптивність розробленого рішення.

Зв'язок із сучасними дослідженнями

Проблематика застосування штучного інтелекту в системах охорони здоров'я активно досліджується. Наприклад, у роботах [1, 2] детально проаналізовано виклики інтеграції алгоритмів штучного інтелекту (ШІ) в медичні системи, включно з оптимізацією обробки великих обсягів даних у реальному часі. Поточні дослідження наголошують на важливості створення гнучких архітектур, здатних інтегрувати нові сервіси для збереження та прогнозування здоров'я пацієнтів. Традиційні медичні системи, як-от електронні медичні записи, часто базуються на монолітній архітектурі, що ускладнює додавання нових функцій і обробку мультимодальних даних.

Основна проблема сучасних методів діагностики апное та респіраторних розладів полягає у залежності від громіздкого обладнання, яке обмежує рухливість пацієнтів і впливає на якість зібраних даних. Наприклад, пристрої для моніторингу сну зазвичай мають стаціонарний дизайн, що знижує комфорт і точність даних, отриманих під час природних рухів. Додатково, різноманітність форматів і відсутність уніфікованих стандартів ускладнюють інтеграцію та обробку отриманих даних. Це створює затримки в діагностиці, підвищує навантаження на медичний персонал і вимагає використання множинних інтерфейсів для аналізу інформації.

Ціллю дослідження є розробка масштабованої системи прогнозування та діагностики респіраторних захворювань, таких як апное сну та хронічне обструктивне захворювання легень, із застосуванням носимих сенсорів. Для цього передбачається дослідити алгоритми фільтрації даних із сенсорів, зокрема фільтр Калмана для зменшення шумів та артефактів; вивчити й оптимізувати моделі штучного інтелекту, зокрема рекурентні нейронні мережі з механізмом «довгої пам'яті», які запам'ятовують тривалі залежності в часових рядах (Long-short term memory (LSTM)), архітектура з механізмом уваги, що одночасно оцінює важливість кожного елемента послідовності (трансформери) та одновірні згорткові мережі, які виявляють локальні патерни в сигналі (One-Dimensional Convolutional Neural Network (1D-CNN)); створити прототипи носимих пристроїв та інтегрувати їх із системою збору й обробки даних; а також розробити хмарну архітектуру для безпечного зберігання, ефективного аналізу та надання результатів клінічним фахівцям.

У ході дослідження була розроблена гнучка мікросервісна архітектура системи, що включає модуль штучного інтелекту (ШІ) для аналізу медичних даних. Носимий пристрій, інтегрований із платформами збору даних, забезпечує безперервний моніторинг стану пацієнтів, мінімізуючи вплив на їхню повсякденну активність. Система надає лікарям доступ до результатів через веб-платформу, де відображаються сигнали фотоплетизмографії після цифрової обробки на носимому пристрої за допомогою фільтра Калмана та їх класифікації рекурентною

нейронною мережею довгої короткочасної пам'яті із трьома шарами по 50 одиниць і методом випадкового відключення нейронів 0,2 для запобігання перенаванчання. Пацієнти через мобільний додаток можуть переглядати історію своїх показників та отримувати структуровані діагностичні висновки у зручному форматі. Для навчання моделі використано публічні набори даних St. Vincent's University Hospital / University College Dublin Sleep Apnea Database, MIT-BIH Polysomnographic Database та Apnea-ECG Database, що дозволило перевірити стійкість алгоритму класифікації на різноманітних сигналах і оптимізувати його точність.

Мета дослідження – створення універсальної та доступної системи моніторингу і класифікації апное під час сну на основі штучного інтелекту, яка забезпечує пацієнтам можливість моніторингу стану дихальної системи за допомогою носимих пристроїв, інтегрованих із мобільними додатками. Система дозволяє збирати, зберігати та аналізувати ключові показники, такі як частота дихання, рівень насичення киснем і серцевий ритм, забезпечуючи пацієнтам доступ до інформації про їхнє здоров'я. Для користувачів це означає надання зворотного зв'язку у формі сповіщень або графіків, що дозволяє вчасно виявляти можливі порушення в роботі дихальної системи. Для лікарів система надає інструменти для збереження та аналізу медичних даних, а також можливість дистанційного моніторингу стану пацієнтів у режимі реального часу через захищені платформи, що підвищує ефективність діагностики та дозволяє виявляти захворювання на ранніх стадіях.

Завдання дослідження:

1. Розробити архітектури системи, яка забезпечує збір, збереження, обробку та прогнозування даних пацієнтів у режимі реального часу. Для цього передбачено використання мікросервісної архітектури, що дозволяє масштабувати систему та легко інтегрувати нові функціональні компоненти.

2. Вивчити та оптимізувати моделі штучного інтелекту — LSTM і 1D-CNN — для аналізу даних, що надходять із носимого датчика MAX30102 (фотоплетизмографія + SpO₂). Опрацьовувані біомедичні сигнали охоплюють

фотоплетизмограму (PPG), насичення крові киснем (SpO_2), частоту серцевих скорочень (ЧСС), частоту дихання та рухову активність грудної клітки. Мета — забезпечити високоточне прогнозування і класифікацію апное сну за рахунок комплексного аналізу цих показників..

3. Дослідити та впровадити алгоритми фільтрації біомедичних даних, зокрема фільтри Кальмана, середньозважені та адаптивні методи, для зменшення шумів і забезпечення високої точності оброблених показників.

4. Розробити та реалізувати механізми захисту передачі даних між носимими пристроями та хмарною системою, зокрема через використання шифрування AES та протоколів TLS/SSL, щоб гарантувати конфіденційність та безпеку інформації про пацієнтів.

Об'єктом дослідження є процеси діагностики та прогнозування респіраторних захворювань, таких як апное сну та хронічне обструктивне захворювання легень, із використанням носимих сенсорних пристроїв та методів штучного інтелекту. Завдання розробки системи для зберігання, обробки та прогнозування даних, а також вивчення моделей штучного інтелекту безпосередньо пов'язані з об'єктом дослідження, що сприяє створенню ефективного інструменту для діагностики та моніторингу захворювань дихальної системи.

Предметом дослідження є застосування штучного інтелекту для аналізу та прогнозування даних і мікросервісної архітектури для обробки та зберігання даних з носимих пристроїв з метою діагностики та прогнозування респіраторних захворювань, а саме класифікації апное сну. Завдання, пов'язані з аналізом моделей ШІ, фільтрацією біомедичних даних і захистом передачі даних, сприяють оптимізації обробки інформації, підвищуючи надійність системи та забезпечуючи мінімальний час затримки між збором і аналізом даних, що в свою чергу гарантує зручність і ефективність використання системи.

У результаті дослідження було створено носимий пристрій, який дозволяє пацієнту використовувати його в умовах повсякденного життя, незалежно від стаціонарних медичних установ. Також розроблено інфраструктуру хмарної

системи, що зберігає, обробляє, прогнозує та надає ці дані лікарям і пацієнтам через веб-платформу або мобільний додаток у зручному для розуміння форматі. Пристрій реєструє серцевий ритм, рівень кисню в крові (SpO_2) і сигнали фотоплетизмографії (PPG), що дозволяє проводити комплексний аналіз дихальної системи, зокрема оцінку частоти дихання та варіабельності ритму серця, на основі даних з одного сенсора.

Для реалізації проєкту було використано методи машинного навчання, зокрема LSTM і 1D-CNN, алгоритми фільтрації біомедичних даних (наприклад, фільтр Кальмана) та технології хмарних обчислень. Ці технології забезпечують збирання, обробку, аналіз і захист даних пацієнтів. Побудова системи здійснювалася на основі мікросервісної архітектури, яка забезпечує гнучкість, масштабованість і можливість інтеграції нових функціональних компонентів без змін основних модулів системи.

Дані отримувалися за допомогою мікроконтролера ESP32-C3, обладнаного сенсором MAX-30102, який реєструє показники фотоплетизмографії (PPG). Ці показники є основою для оцінки серцевого ритму, рівня кисню в крові (SpO_2) та інших параметрів дихальної системи. Зв'язок між пристроєм та мобільним телефоном здійснювався за допомогою протоколу BLE (англ. Bluetooth Low Energy), що забезпечує стабільну передачу даних із низьким енергоспоживанням для збільшення автономності носимого пристрою. Передача даних до хмарного сервісу здійснювалася через протокол MQTT, який завдяки мінімальним накладним витратам і підтримці публікаційно-підписної моделі передачі даних забезпечує легковагу та ефективність зв'язку. Це особливо важливо для роботи з великими обсягами сенсорних даних.

Для підвищення якості даних і зниження шумів у показниках використовувався фільтр Калмана, реалізований безпосередньо на мікроконтролері. Цей метод дозволяє усувати низькочастотного дрейфу ізолінії ($< 0,5$ Гц), мережевих наведень 50 Гц (60 Гц) і рухових артефактів у смузі 1–15 Гц — застосовано каскад високопрохідного фільтра (0,5 Гц) та вузькосмугового смуго загороджувального (англ. notch) фільтра, після яких працює адаптивний фільтр

Калмана, реалізований безпосередньо на мікроконтролері. З метою обробки сигналів ФПГ реалізовано алгоритми, що аналізують амплітуду та частотний склад сигналів для обчислення серцевого ритму та рівня кисню в крові (SpO_2). Ці показники є критично важливими для забезпечення високої точності початкової інформації, яка надалі використовується моделлю ШІ для прогнозування стану пацієнта.

Для аналізу біомедичних показників і прогнозування можливих респіраторних захворювань використовувалася модель довготривалої короткострокової пам'яті (LSTM). Цей тип рекурентної нейронної мережі був обраний завдяки здатності зберігати та обробляти контекстні залежності у часових рядах, що є ключовим для врахування динамічних змін у даних PPG. Модель LSTM використовує механізми комірок пам'яті, що дозволяють зберігати важливу інформацію та відсіювати несуттєві дані у часових рядах. Це забезпечує можливість точного прогнозування показників дихальної системи, зокрема частоти дихання, рівня кисню в крові (SpO_2) та варіабельності серцевого ритму.

Для оцінки якості моделі використовувалася функція втрат Mean Squared Error (MSE) на етапі навчання, що дозволило визначити ефективність моделі у процесі оптимізації. Додатково оцінка проводилася шляхом порівняння прогнозованих значень із реальними даними тестового набору, що забезпечило глибший аналіз точності прогнозування. Додаткові метрики не використовувалися на початковому етапі дослідження, оскільки основна увага зосереджувалася на базовій працездатності моделі та відповідності її прогнозів реальним даним.

Передача даних між сенсором і мобільним пристроєм здійснювалася за допомогою шифрування за алгоритмом AES-256 із використанням унікального ключа, встановленого на кожен пристрій за замовчуванням. Це забезпечує захист від несанкціонованого доступу до даних під час їх передачі. Для захисту даних, що передаються до хмарного сервісу, було реалізовано авторизацію та аутентифікацію з використанням токенів JWT, які підписуються за допомогою алгоритму HMAC або RSA, що гарантує цілісність і автентичність інформації. Передача

здійснювалася через захищений протокол HTTPS, який забезпечує додатковий рівень безпеки під час роботи з чутливими медичними даними пацієнтів.

Архітектура системи розроблена на основі мікросервісної структури із використанням кінцевих обчислень (edge computing) та подієвого дизайну (event-driven architecture). Edge computing забезпечує обробку даних безпосередньо на носимому пристрої, що мінімізує затримки та оптимізує використання хмарних ресурсів. Взаємодія між сервісами здійснюється через API-шлюз (наприклад, Kong або AWS API Gateway), який централізує управління запитами між сервісами, забезпечує модульність і спрощує масштабування кожного компонента. Кожна функціональна частина системи реалізована як окремий сервіс, що дозволяє оновлювати або додавати нові функціональні компоненти без зупинки чи перезапуску інших сервісів, забезпечуючи гнучкість і адаптивність системи.

Сукупність зазначених методів забезпечує ефективну обробку біомедичних даних у режимі реального часу, зберігаючи високу точність прогнозування, надійність роботи алгоритмів та захищеність системи. Це є ключовим для своєчасного виявлення та прогнозування респіраторних захворювань, що сприяє підвищенню якості моніторингу та діагностики стану пацієнтів.

Зв'язок роботи з науковими програмами, планами, темами. Основний зміст роботи складають результати досліджень, які проводились протягом 2021–2025 років, відповідно до наукового напрямку кафедри «Системного проектування» Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» та держбюджетної теми №18.2021.ММ «Інтелектуальний сервіс діагностики стану дихальної системи людини для зниження ризиків в умовах пандемії COVID-19» (№ ДР 0121U110623), в якій автор був співвиконавцем.

Методи дослідження. В основу досліджень покладено методи сучасної цифрової обробки біомедичних сигналів та штучного інтелекту, такі як: адаптивний фільтр Кальмана; моделювання послідовностей із використанням LSTM; прикінцеві (edge) обчислення на носимому пристрої; мікросервісна

архітектура для масштабованої обробки й аналізу даних; а також методи експериментальних випробувань із прототипом носимого сенсорного модуля.

Наукова новизна отриманих результатів

1. Вперше розроблено метод раннього прогнозування респіраторних станів та поглибленої діагностики апное сну, який, на відміну від існуючих підходів, що обмежуються переважно моніторингом або загальним виявленням апное, забезпечує:

- о Детальну класифікацію типів апное сну (обструктивного, центрального та змішаного) на основі інтегрованого аналізу даних фотоплетизмографії (ФПГ) та персоніфікованих статичних показників пацієнта за допомогою запропонованої моделі штучного інтелекту.

- о Оцінку динаміки індексу апное-гіпопное (АHI), що дозволяє відстежувати зміни стану пацієнта в часі та ефективність превентивних або терапевтичних заходів.

- о Проактивне прогнозування можливих ускладнень та погіршень стану апное, надаючи користувачеві своєчасні попередження для вжиття превентивних заходів.

2. Обґрунтовано та реалізовано інноваційний підхід до обробки біомедичних даних для моніторингу дихання та апное із застосуванням прикінцевих обчислень (edge computing) безпосередньо на смартфоні користувача. Це вперше дозволяє для даного класу задач суттєво мінімізувати обсяг передачі чутливих медичних даних у хмару, знизити енергоспоживання носимих пристроїв та смартфона, забезпечити тривалу автономну роботу системи та підвищити її надійність в умовах обмеженого доступу до мережі Інтернет.

3. Запропоновано оригінальну гнучку, масштабовану та безпечну мікро сервісну архітектуру мобільної телемедичної системи. На відміну від монолітних або менш адаптивних систем, розроблена архітектура через стандартизовані API дозволяє:

- о Легко інтегрувати нові типи біомедичних сенсорів.
- о Оперативно впроваджувати та оновлювати моделі штучного інтелекту.

- о Забезпечувати універсальність системи шляхом її швидкої адаптації для моніторингу та прогнозування інших типів захворювань через заміну відповідних сенсорних модулів та алгоритмів ШІ без необхідності кардинальної перебудови базової інфраструктури.

- о Гарантувати високий рівень захисту даних на всіх етапах їх обробки та передачі за допомогою сучасних криптографічних методів (наприклад, AES-256) в рамках розподіленої архітектури.

4. Науково обґрунтовано синергетичний ефект від поєднання розроблених методів предиктивної діагностики, архітектури прикінцевих обчислень та гнучкого мікросервісного підходу, що вперше уможливило створення практично орієнтованої, зручної для користувача мобільної системи. Ця система забезпечує не лише безперервний та довготривалий моніторинг стану дихальної системи пацієнта в умовах повсякденного життя, але й надає інструменти для проактивного управління здоров'ям, що суттєво відрізняє її від стаціонарних клінічних методів та існуючих мобільних аналогів, орієнтованих переважно на обмежений моніторинг.

На відміну від інших досліджень, зосереджених на вузьких аспектах діагностики дихальних розладів, таких як аналіз кашлю чи дослідження показників легенів [11, 12], розроблена система інтегрує можливості фотоплетизмографії (ФППГ) з прогнозуванням респіраторних захворювань. Більшість існуючих робіт із PPG обмежуються моніторингом показників без спроб передбачення захворювань [13]. Розроблена мобільна телемедична система на базі носимого датчика PPG + акселерометра та мікросервісної хмарної платформи розширює межі діагностики, поєднуючи безперервний моніторинг і прогнозування. Це дає змогу здійснювати раннє виявлення та класифікацію апное сну пацієнтів на основі біомедичних показників, таких як насичення крові киснем (SpO_2), частота серцевих скорочень (ЧСС), частота дихання, індекс апное-гіпопное (АHI), рівень рухової активності, а також статичних показників статі, віку та ваги. Гнучка та масштабована архітектура є ще однією інноваційною складовою системи. Побудована на основі мікро сервісного підходу, вона дозволяє інтегрувати нові сенсори та моделі

штучного інтелекту через стандартизовані API без значних змін в основній інфраструктурі. Це забезпечує універсальність, даючи можливість адаптувати систему до діагностики інших захворювань шляхом простої заміни сенсорів або навчання нових моделей ШІ. Система орієнтована на забезпечення зручності для користувачів завдяки мобільності, простоті використання та можливості моніторингу поза межами клінічних умов. Більшість стаціонарних методів моніторингу є складними та недоступними для тривалого застосування в повсякденному житті. Натомість запропонована система є вигідною альтернативою для довготривалого моніторингу здоров'я вдома. Завдяки можливості постійного контролю показників дихальної системи пацієнти можуть вчасно виявляти потенційні проблеми зі здоров'ям і звертатися за медичною допомогою, що значно покращує прогноз і якість життя.

Архітектура системи базується на перевірених сучасних підходах, що поєднують гнучкість і безпеку мікросервісних структур. Кожен компонент виконує специфічну роль, наприклад, збирання даних, їх обробка, зберігання або управління доступом. Такий підхід забезпечує легкість масштабування, адаптацію до нових вимог і інтеграцію з іншими компонентами. Для взаємодії між компонентами використовуються унікальні ключі шифрування, зокрема алгоритми AES-256, які гарантують надійний захист даних та конфіденційність. Завдяки цьому система відповідає сучасним стандартам інформаційної безпеки, забезпечуючи безпеку навіть у сценаріях із високими ризиками.

Основна відмінність розробленої системи від існуючих полягає в акценті на прогнозуванні, а не лише на моніторингу. Система не лише вимірює показники дихання, а й прогнозує можливі проблеми, надаючи користувачам попередження. Це дозволяє отримувати своєчасну інформацію про стан здоров'я, приймати превентивні заходи та уникати можливих ускладнень. Новизна роботи полягає в інтеграції прогнозування на основі біомедичних даних, гнучкої архітектури та орієнтованості на практичне використання в умовах повсякденного життя. Завдяки поєднанню цих характеристик система не лише здійснює безперервний моніторинг, але й забезпечує прогноз і класифікацію апное сну — обструктивного,

центрального та змішаного типів — з оцінкою динаміки індексу апное-гіпноє (АHI). Такий підхід підвищує ефективність довготривалого спостереження за пацієнтами з порушеннями дихання під час сну.

Практичне значення отриманих результатів полягає у наступному:

Особистий внесок здобувача.

Усі теоретичні та практичні результати дисертаційної роботи, що виносяться на захист, отримані здобувачем особисто. У друкованих працях, написаних у співавторстві, здобувачеві належать:

- У статті «Особливості застосування алгоритму EMD для оцінки рівня кисню в крові» проведено аналіз сигналів пульсоксиметрії й застосовано емпіричне модове розкладання для виявлення десатурацій. Здобувач запропонував і реалізував детектор десатурації на основі вибраних меж та набору правил, побудував індекс десатурації кисню для діагностики і класифікації апное сну та оцінив ефективність підходу (чутливість = 0,838).
- У статті «Узагальнення інформації з прикладами про можливості використання сервіс-орієнтованого підходу та технологій штучного інтелекту в e-health» здобувач дослідив способи упровадження SOA та AI у сучасні медичні системи. Здобувач систематизував існуючі архітектури, узагальнив переваги інтеграції сервісів та штучного інтелекту і сформував рекомендації щодо їх застосування для класифікації апное сну.
- У статті «Development of a patient health monitoring system based on a service-oriented architecture using artificial intelligence» проаналізовано мікросервісну систему моніторингу пацієнтів, що поєднує SOA й AI. Здобувач розробив архітектуру, інтегрував дані з різних джерел, упровадив LSTM-модуль класифікації апное сну й продемонстрував переваги запропонованого рішення щодо масштабованості та точності.
- У статті «Diagnosis of respiratory rhythm based on physiological parameters of plethysmography using artificial intelligence» досліджено застосування LSTM-моделі для розрахунку частоти дихання на основі PPG-сигналів, віку та статі пацієнтів. Здобувач сформував датасет, навчив і протестував модель,

показавши придатність підходу для підтримки класифікації апное сну та відмітивши залежність точності від вікових груп.

- У статті «Розробка інфраструктури для інтеграції штучного інтелекту в телемедицинні системи за допомогою мікросервісної архітектури» представлено мікросервісну архітектуру з message-broker-ом та службами HL7 FHIR. Здобувач спроектував інфраструктуру, реалізував автоматизований pipeline навчання й деплоюменту моделей класифікації апное сну та обґрунтував переваги рішення для реального часу.
- У статті «Дослідження ефективності алгоритмів фільтрації біомедичних сигналів у реальному часі» оцінено низькочастотні, ковзні, середньозважені та Калман-фільтри для PPG-сигналів, зібраних на ESP32 + MAX30102. Здобувач реалізував усі алгоритми, зібрав та проаналізував дані, встановив перевагу Kalman-фільтра й довів, що його застосування підвищує якість сигналу та точність модулів класифікації апное сну у домашньому моніторингу.

Апробація результатів дисертації.

Публікації.

1. Boloban, O., Pysmennyi, I., Kyslyi, R., & Kyriusha, B. (2024). Development of a patient health monitoring system based on a service-oriented architecture using artificial intelligence. *Technology Audit and Production Reserves*, 3(2(77)), P. 23–29. (Фахове видання категорії “Б”) (Здобувачу належить реалізація програмної частини, порівняльний аналіз запропонованих методів)
2. Петренко, А. І., Болобан, О. А. (2025). Розробка інфраструктури для інтеграції штучного інтелекту в телемедицинні системи за допомогою мікросервісної архітектури. *Таврійський науковий вісник. Серія: Технічні науки*, (1), С. 148-158. (Фахове видання категорії “Б”) (Здобувачу належить аналіз літературних джерел, аналіз та порівняння існуючих підходів та методів, реалізація програмної частини, обґрунтування та узагальнення отриманих результатів)

3. Болобан О.А., Петренко А.І. Особливості застосування алгоритму EMD для оцінки рівня кисню в крові. Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки. Том 34 (73). № 2. 2023. С. 125-130. (Фахове видання категорії “Б”) (Здобувачу належить постановка задачі, аналіз літературних джерел, аналіз та порівняння існуючих підходів та методів)
4. Петренко, А., Болобан, О. (2023). Узагальнення інформації з прикладами про можливості використання сервіс-орієнтованого підходу та технологій штучного інтелекту в галузі e-health. Технологічний аудит та резерви виробництва, 4 (2 (72)), С. 10-17. (Фахове видання категорії “Б”) (Здобувачу належить постановка задачі, аналіз наявних рішень, реалізація програмної частини, обґрунтування та узагальнення отриманих результатів)
5. Петренко, А., Болобан, О. (2024). Дослідження ефективності алгоритмів фільтрації біомедичних сигналів у реальному часі. Інфокомунікаційні та комп’ютерні технології, 2(08), С. 53-61. (Фахове видання категорії “Б”) (Здобувачу належить аналіз літературних джерел, аналіз та порівняння існуючих підходів та методів, реалізація програмної частини, обґрунтування та узагальнення отриманих результатів)

Структура дисертації та її обсяг.

Дисертація складається із вступу, шести розділів, двадцяти трьох підрозділів, загальних висновків, списку використаних джерел та додатків. Загальний обсяг роботи складає 156 сторінок, із яких 124 сторінки основного тексту. У роботі міститься 43 рисунки, 10 таблиці. Список використаних джерел налічує 69 найменування, із яких 69 іноземною мовою.

РОЗДІЛ 1. ВПРОВАДЖЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ В ОХОРОНУ ЗДОРОВ'Я: МОЖЛИВОСТІ ТА ВИКЛИКИ

Згідно з Леманном [3], штучний інтелект (ШІ) трансформував традиційні ручні методи в автоматизовані системи, здатні аналізувати великі обсяги медичних даних і підтримувати прийняття клінічних рішень. Системи на основі ШІ виконують діагностичні та прогностичні задачі швидше, точніше та ефективніше, ніж це можливо людиною [4]. Разом із тим, інтеграція ШІ в охорону здоров'я потребує врахування етичних аспектів (конфіденційність, відповідальність) та технічних викликів (несумісність ІТ-систем, упередженість даних) [5–8].

Водночас останні дослідження зосереджені на використанні методів машинного і глибинного навчання для класифікації апное сну на основі біомедичних сигналів із носимих пристроїв. Так, модель ApSense, що аналізує лише PPG-сигнали, поєднує домен-специфічні ознаки (амплітуда хвилі, інтервали RR тощо) та глибинну нейронну мережу, і демонструє вищу чутливість та AUROC порівняно з SOTA-підходами при виявленні епізодів обструктивного апное сну. Інші роботи показали, що поєднання селекції ознак і класичних ML-алгоритмів (SVM, Random Forest, XGBoost) дозволяє прогнозувати ступінь тяжкості ОАС з точністю понад 82 %.

Разом із тим, впровадження ШІ в медицину супроводжується етичними питаннями та занепокоєннями, зокрема щодо конфіденційності даних і відповідальності за прийняті рішення, що впливає на довіру до цих технологій. Використання алгоритмів глибокого навчання для діагностики раку, прецизійної медицини, прогнозування та автоматизації управлінських завдань стикається з викликами, зокрема складнощами забезпечення надійності моделей ШІ, їх адаптації до різних клінічних умов та типів даних, а також балансування між машинними рішеннями і людським досвідом у прийнятті рішень [5].

Субасі [6] зазначає, що однією з найбільших переваг використання штучного інтелекту (ШІ) в медицині є підвищення точності діагностики та загальної ефективності медичних процедур. Інструменти візуалізації на основі ШІ, такі як згорткові нейронні мережі, здатні аналізувати медичні зображення, зокрема

магнітно-резонансну томографію (МРТ) і комп'ютерну томографію (КТ), з високою точністю, перевершуючи людський аналіз.

Завдяки здатності ІІІ виявляти дрібні аномалії, які можуть бути непомітними для людського ока, ці технології дозволяють проводити раннє виявлення багатьох захворювань, зокрема раку. Раннє виявлення значно покращує прогноз для пацієнтів, дозволяючи розпочати лікування на початкових стадіях хвороби та підвищуючи шанси на успішне одужання.

Щодо використання штучного інтелекту у медичній сфері, Шіннерс та ін. [7] досліджували й зазначили, що, попри значний потенціал, інтеграція ІІІ в медичні системи супроводжується технічними, етичними та організаційними викликами. Однією з основних проблем є несумісність між системами ІІІ та існуючими ІТ-системами в медицині. Більшість сучасних систем охорони здоров'я працюють на застарілому програмному забезпеченні, яке не підтримує обробку великих обсягів даних або алгоритми машинного навчання, що ускладнює інтеграцію з технологіями ІІІ, а також збільшує вартість модернізації.

Крім того, Шіннерс та ін. [7] наголошують, що питання конфіденційності та безпеки даних є ключовими у медичній сфері. ІІІ може підтримувати системи охорони здоров'я в обробці великих обсягів даних пацієнтів, забезпечуючи високий рівень безпеки та конфіденційності інформації. Організації, які впроваджують ІІІ, отримують інструменти для підвищення захисту конфіденційних даних, зокрема через шифрування та автоматизовану ідентифікацію потенційних загроз.

Інше питання, пов'язане із ІІІ, як зазначають Суджіт та ін. [8], полягає у можливості поглиблення існуючих нерівностей у сфері охорони здоров'я. Алгоритми ІІІ є настільки якісними, наскільки якісні дані, на яких вони були навчені. Наприклад, якщо дані містять недостатню репрезентацію певних груп населення за расовою, гендерною чи соціально-економічною ознакою, це може спричинити упередженість у прогнозах або рекомендаціях ІІІ, що призводить до нерівного лікування пацієнтів. Аналогічно, Субасі [6] підкреслює етичні виклики, пов'язані з роллю людського судження у порівнянні з машинними рішеннями. Хоча

ІІІ може значно підтримувати процеси прийняття рішень, такі важливі аспекти, як інтуїція, емпатія та сприйняття, не можуть бути інтегровані у ці системи. Складні або неоднозначні медичні рішення часто потребують інтеграції людського досвіду та індивідуального підходу до пацієнта, що залишається поза межами можливостей ІІІ.

Згідно з висновками Янга та ін. [9], ІІІ має потенціал запропонувати більш пацієнт-орієнтований і безперервний догляд, зокрема через моніторинг стану пацієнтів у реальному часі та адаптацію лікувальних планів до їхніх індивідуальних потреб. Системи ІІІ, використовуючи дані з носимих пристроїв, забезпечують медичних працівників інформацією про поточний стан хронічних пацієнтів, дозволяючи аналізувати динаміку та розробляти індивідуалізовані плани лікування.

Захворювання, такі як діабет, гіпертензія та серцева недостатність, можуть бути краще контрольовані завдяки використанню носимих пристроїв, що забезпечують постійний моніторинг і допомагають уникати ускладнень. ІІІ також сприяє дотриманню лікувальних режимів пацієнтами через функції нагадувань, трекерів прийому ліків і автоматизованих повідомлень. Це особливо корисно у складних випадках із багатокомпонентними схемами лікування, зменшуючи ризики пропусків або неправильного прийому препаратів.

Точність діагностичних інструментів ІІІ на 15% перевищує традиційні методи при виявленні ранніх стадій захворювань. Це також скорочує час діагностики на 30%, що прискорює процес тріажу та лікування пацієнтів. У сфері персоналізованої медицини ІІІ демонструє на 20% вищу точність у передбаченні реакції пацієнтів на лікування, що значно покращує результати терапії. Крім того, 40% опитаних медичних працівників повідомили про зменшення адміністративного навантаження завдяки використанню ІІІ, що звільнило більше часу для безпосередньої роботи з пацієнтами [10]. Кейс-стадії також показали, що впровадження ІІІ скорочує витрати на медичне обслуговування на 25% завдяки оптимізації розподілу медичного персоналу, ресурсів лікарень та автоматизації адміністративних процесів. Попри значний позитивний вплив ІІІ, результати

також виявили низку викликів. Точність алгоритмів була на 10% нижчою при діагностиці представників певних етнічних та соціальних груп, що вказує на недостатню репрезентативність таких пацієнтів у навчальних вибірках [7]. Це підкреслює необхідність використання більш інклюзивних даних для навчання ШІ. Статистичні аналізи підтвердили значний зв'язок між використанням ШІ та покращенням результатів лікування пацієнтів (значення $p < 0,05$), що підтверджує гіпотезу про позитивний вплив ШІ на медицину. Однак 60% медичних працівників висловили побоювання щодо надмірної залежності від ШІ, і 60,3% наголосили на необхідності інтеграції між людським та штучним інтелектом. Вони вважають, що ШІ має слугувати інструментом підтримки людського інтелекту, а не змінювати його.

РОЗДІЛ 2. МЕТОДОЛОГІЧНИЙ АНАЛІЗ ПРОБЛЕМИ ДІАГНОСТИКИ АПНОЕ СНУ З ВИКОРИСТАННЯМ ГЛИБОКОГО НАВЧАННЯ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ДІАГНОСТИЧНИХ РІШЕНЬ ЛІКАРЕМ

2.1 Захворювання органів дихання та причини порушення дихання та виникнення апное сну

Патології органів дихання охоплюють значний спектр захворювань, що порушують вентиляцію дихальних шляхів. Найбільш розповсюдженими серед них є хронічна обструктивна хвороба легень, астма бронхіальна та інтерстиціальні ураження легеневої тканини, які призводять до змін еластичності паренхіми та звуження бронхіального просвіту. Підґрунтям цих порушень найчастіше є не лише запальні чи інфекційні процеси, а й структурні аномалії дихальних шляхів, дисбаланс імунних реакцій або токсичний вплив шкідливих речовин. Крім того, на ефективність дихання впливають інші фактори — ожиріння, куріння, алергічні реакції та екологічні чинники, які ускладнюють роботу дихальних м'язів і знижують здатність легенів до адекватної вентиляції.

Апное сну, як особлива форма порушення дихання, виникає під час сну через періодичні зупинки повітрообміну тривалістю не менше десяти секунд. Обструктивне апное сну зумовлено спаданням стінок глотки внаслідок зниження

тону м'язів верхніх дихальних шляхів, що призводить до повної або часткової блокади повітряного потоку. Ключовими факторами ризику є анатомічні особливості черепно-лицевої ділянки, надмірна вага тіла та вік, адже у людей середнього та старшого віку тонус м'язів глотки під час сну значно знижується. Центральне апное, навпаки, виникає внаслідок недостатньої активації дихального центру в головному мозку, коли пацієнт фізично не здійснює дихальних рухів, хоча прохідність дихальних шляхів залишається незмінною. У багатьох випадках зустрічається змішана форма, коли центральні і обструктивні механізми поєднуються, ускладнюючи клінічну картину патології.

Патофізіологічні наслідки апное сну полягають передусім у хронічній гіпоксії та частих реактиваціях симпатичної нервової системи, що провокують стрибки артеріального тиску, порушення серцевого ритму і фрагментацію сну. Це зумовлює розвиток артеріальної гіпертензії, метаболічних розладів, посилює ризик серцево-судинних ускладнень та знижує якість життя через постійну денну сонливість і когнітивні дисфункції. Урахування цих чинників є визначальним для вибору ефективних методів діагностики та лікування: від консервативних підходів із застосуванням безперервного позитивного тиску в дихальних шляхах (CPAP) до хірургічних втручань, спрямованих на корекцію анатомічних порушень. Саме тому розробка автоматизованих систем моніторингу та прогнозування апное сну на основі біомедичних сигналів і сучасних методів штучного інтелекту набуває особливого значення в контексті покращення результатів лікування та профілактики ускладнень.

2.2 Традиційні клінічні методи реєстрації та аналізу біомедичних сигналів для діагностики апное сну

Традиційні клінічні методи діагностики апное сну ґрунтуються насамперед на багатоканальній полісомнографії, яка одночасно реєструє електроенцефалограму (ЕЕГ), електроміограму (ЕМГ), електроокулограму (ЕОГ), дихальний потік через назально-ротовий термінал та термоструминний сенсор, перфузію крові й насичення киснем (SpO_2) за допомогою пульсоксиметра, а також рух грудної клітки й черевної стінки за допомогою оптичних або тензометричних

ременів (рис. 1.1.). Паралельно фіксується серцевий ритм по електрокардіограмі (ЕКГ) та положення тіла із застосуванням акселерометричних матриць.



Рис. 1.1. Полісомнографічна система

Отримані сигнали аналізуються як автоматизованими програмно-апаратними комплексами, так і вручну — шляхом перегляду фрагментів зупинок дихання, десатурацій і пробуджень, що дозволяє визначити індекс апное-гіпноє (АHI) та ступінь тяжкості порушення (легкий, середній, важкий). Для оцінки дихального потоку застосовують каналні фільтри та кореляційний аналіз між рухами грудної клітки й змінною температурою повітря в каналі дихання. Сигнали SpO_2 обробляються за допомогою лінійної інтерполяції та відстеження піків деструкцій, що дає змогу обчислити час та глибину гіпоксичних епізодів (рис. 1.2.).



рис. 1.2. Пульсоксиметр

Незважаючи на високу інформативність, полісомнографія потребує складного обладнання, залучення кількох операторів і може спричиняти дискомфорт пацієнту через численні датчики та дроти. Аналіз даних часто займає кілька годин і вимагає підтвердження експертом-сомнологом, що обмежує масштабованість методу для масових скринінгів. Саме ці недоліки стали каталізатором пошуку менш інвазивних підходів із носимими пристроями та автоматизованими алгоритмами обробки сигналів.

2.3 Застосування методів цифрової обробки сигналів для покращення якості діагностики апное сну

У сучасних дослідженнях діагностики апное сну дедалі частіше застосовують комплекс методів цифрової обробки сигналів, що дозволяють суттєво підвищити точність і швидкість виявлення індивідуальних епізодів зупинок дихання. На етапі попередньої обробки сигнали пульсоксиметрії і кардіосигнали проходять через низькочастотні високопрохідні фільтри для усунення дрейфу ізоїнії й смугозагороджувальні (англ. notch) фільтри для придушення мережових наведень 50/60 Гц. Для адаптивного виділення корисної пульсової хвилі часто використовують фільтр Кальмана або незалежний компонентний аналіз, які

зберігають ключові часово-амплітудні характеристики сигналу та водночас зменшують артефакти руху й шум пристрою. Після очищення застосовують часово-частотні перетворення: швидке перетворення Фур'є виділяє характерні частотні компоненти дихання (приблизно 0,1–0,5 Гц), а вейвлет-аналіз локалізує короточасні гіпоксичні епізоди в часовій області [59]. З отриманих часово-частотних ознак вручну або з використанням простих правил детектування (thresholding, envelope detection, пошук пікових інтервалів RR) формують події зупинок дихання та розраховують індекс апное-гіпопное (AHI), який є основним клінічним показником тяжкості порушення сну.

Ключовим кроком є також верифікація правильності сегментації та класифікації апное шляхом аналізу еволюції спектра і статистичних характеристик сигналу—наприклад, зміни варіабельності інтервалів R-R або амплітудні модулюючі ефекти у PPG під час дихальних пауз. Такий підхід, заснований винятково на класичних алгоритмах цифрової обробки сигналів, підвищує точність і надійність діагностики апное сну без необхідності залучати експертів для ручного перегляду всіх записів.

2.4 Аналіз застосування методів глибокого навчання до задачі класифікації сигналів фотоплетизмограм для виявлення апное сну

Упродовж останніх років застосування методів глибинного навчання до класифікації сигналів фотоплетизмографії для виявлення апное сну демонструє значний прогрес у підвищенні точності діагностики. Низка досліджень показала, що одновимірні згорткові нейронні мережі (англ. One-Dimensional Convolutional Neural Network (1D-CNN)) можуть автоматично виділяти часові ознаки без попереднього ручного інженірингу, досягаючи точності виявлення понад 90% при розпізнаванні епізодів пауз дихання. Зокрема, у роботі, опублікованій у *Frontiers in Neuroscience*, модель на основі 1D-CNN продемонструвала перевищення 90% точності на тестовому наборі даних PPG, підтверджуючи життєздатність портативних носимих пристроїв для скринінгу апное сну [60].

Поряд із CNN, глибокі рекурентні мережі LSTM та їх гібридні комбінації з CNN (CNN-LSTM) успішно застосовуються для обробки послідовностей ФПП із

врахуванням довготривалих залежностей сигналу. Так, гібридна модель CNN-LSTM досягла 95,38 % точності виявлення апное, підкреслюючи важливість інтеграції просторового й часово-дінамічного аналізу ФПГ-даних [61]. Крім того, платформа OxiNet, розроблена й описана в журналі Nature Communications, продемонструвала високу стійкість до різних розподілів даних та перевершила класичні підходи машинного навчання в діагностиці обструктивного апное за даними пульсоксиметрії [62].

Інші дослідження, зібрані в систематичних оглядах, підтверджують, що моделі на основі CNN, LSTM або їх комбінацій стабільно перевищують 80–90% у чутливості та специфічності класифікації апное сну, особливо коли вхідними даними є часово-частотні представлення ФПГ-сигналу [63]. Такі моделі дозволяють реалізовувати повністю автоматизовані конвеєри діагностики, зменшуючи потребу у ручному аналізі полісомнографії та відкриваючи шлях до масштабного домашнього моніторингу.

Завдяки здатності глибоких мереж опановувати складні патерни в часово-амплітудних даних без попередньої екстракції ознак, ці підходи стають ключовими компонентами сучасних носимих систем для класифікації апное сну. У наступних розділах буде детально розглянуто архітектури LSTM і 1D-CNN, їхню реалізацію у вигляді мікро сервісних компонентів та результати апробації на експериментальних даних.

2.4.1. Виділення репрезентативних ознак, які можуть бути пов'язані з апное сну

Нормальне дихання залежить від кількох факторів, зокрема віку, фізичної підготовки та рівня стресу, і зазвичай характеризується стабільною частотою та обсягом. У новонароджених частота дихання становить від 30 до 60 вдихів за хвилину, тоді як у дорослих цей показник зазвичай коливається в межах 12–20 вдихів за хвилину. Частота дихання може зростати внаслідок фізичної активності, стресу або захворювань. У розслабленому стані або при використанні дихальних технік, наприклад, під час медитації, деякі люди здатні знизити частоту дихання до

трьох-п'яти вдихів за хвилину, що сприяє покращенню самопочуття та зниженню рівня стресу.

Таб. 2.1. Діапазон частоти дихання [16]

| Група | Вік | Дихальних рухів/хв |
|---------------------|----------------------------|--------------------|
| Новонароджені | від народження до 6 тижнів | 30–60 |
| Немовлята | від 6 тижнів до 6 місяців | 25–40 |
| Діти раннього віку | від 1 до 3 років | 20–30 |
| Діти молодшого віку | від 3 до 6 років | 20–25 |
| Діти старшого віку | від 10 до 14 років | 15–20 |
| Дорослі | дорослі | 12–20 |

У пацієнтів з підозрою на апное сну ключовими репрезентативними ознаками (ознаками) стану є ті параметри, які прямо відображають порушення вентиляції та оксигенації під час сну. По-перше, це індекс апное-гіпопное (АHI) – середня кількість подій повної зупинки дихання (апное) та часткових обмежень потоку повітря (гіпопное) за годину сну. За значенням АHI класифікують тяжкість захворювання (таблиця 2.2). По-друге, використовують оксигенаційний індекс (ODI) – кількість десатурацій (падінь SpO_2 на $\geq 3\%$) за годину; а також середній і мінімальний рівні насичення крові киснем (Mean SpO_2 , Min SpO_2). По-третє, фіксують тривалість кожної апное- та гіпопное-події, їхню варіабельність і співвідношення часу пауз до загального сну – це дозволяє оцінити ступінь порушення вентиляції.

Таб. 2.2. Індекс апное-гіпопное (АHI), подій/год

| Індекс апное-гіпопное (АHI), подій/год | Класифікація | Опис |
|--|--------------|--|
| АHI < 5 | Норма | Відсутність клінічно значущих пауз дихання під час сну |

| | | |
|---------------------------|---------------|--|
| $5 \leq \text{АHI} < 15$ | Легка форма | Легкі порушення сну: від 5 до 15 подій апное-гіпноное на годину сну |
| $15 \leq \text{АHI} < 30$ | Середня форма | Помірні порушення сну: від 15 до 30 подій апное-гіпноное на годину сну |
| $\text{АHI} \geq 30$ | Важка форма | Важкі порушення сну: 30 і більше подій апное-гіпноное на годину сну |

Виділяють три основні типи апное сну. Обструктивне апное (ОАС) виникає через спадання стінок глотки внаслідок зниження тону м'язів, що призводить до механічної блокади дихальних шляхів. Центральне апное (ЦАС) зумовлене недостатньою генерацією команд дихальним центром головного мозку – пацієнт не робить спроб вдиху, хоча просвіт дихальних шляхів залишається відкритим. Змішане апное поєднує характеристики обох форм: зазвичай спочатку настає центральна пауза, а потім долучається обструктивний компонент. Кожен із цих станів несе власні ризики. ОАС асоційоване із розвитком артеріальної гіпертензії, серцево-судинних подій, метаболічними порушеннями та денним уповільненням реакцій через фрагментацію сну. ЦАС найчастіше трапляється при тяжкій серцевій недостатності чи неврологічних захворюваннях, підвищуючи ризик серцевих аритмій та інсультів. Змішане апное поєднує обидва механізми та часто має найгірший прогноз.

У сучасних сомнологічних лабораторіях діагностика апное сну базується на полісомнографії — багатоканальній реєстрації ЕЕГ, ЕМГ, ЕОГ, потоків повітря через термістори або пневмотахографи, ременів дихальних зусиль на грудях і животі, параметрів SpO_2 та ЕКГ. Записи аналізуються автоматизованими системами та вручну експертами-сомнологами, що вимагає складної апаратури, участі кількох спеціалістів і може бути незручно для пацієнта.

Мобільна охорона здоров'я (m-health) є однією з перспективних областей моніторингу пацієнтів, яка отримала значну увагу в останні роки [17, 18]. Однак пацієнти вдома часто уникають використання адгезивних датчиків, таких як електроди ЕКГ, через дискомфорт і незручність їх тривалого носіння [19, 20]. У

зв'язку з цим, m-health додатки переважно інтегрують менш інвазивні методи моніторингу, наприклад, пульсоксиметрію. Цей метод дозволяє пацієнтам легко вимірювати рівень насичення киснем крові, просто вставляючи палець у датчик пульсоксиметра, що значно підвищує комфортність використання технологій у домашніх умовах.

Ринок споживчої електроніки в даний час переповнений великою кількістю фітнес-трекерів та інших подібних пристроїв. Основна мета таких систем полягає не у виявленні фізіологічного погіршення, а в підтримці та оптимізації "добробуту" здорових осіб. Це є особливо привабливим для виробників споживчої електроніки, оскільки такі додатки не вимагають дорогого та тривалого клінічного підтвердження, необхідного для сертифікації пристроїв для медичного використання. Більшість існуючих пристроїв оснащені простими акселерометрами для відстеження фізичної активності. Проте останнє покоління цих пристроїв включає більш складні сенсори, такі як пульсоксиметри (наприклад, у нових моделях розумних годинників Apple, Inc., США) та біоімпедансні сенсори (наприклад, у пристроях UP3 від Jawbone, США). Кардіосинхронізовані сигнали, отримані з таких пристроїв, наразі використовуються переважно для визначення частоти серцевих скорочень (ЧСС, англ. heart rate (HR)). Водночас, за умови доступності достатньо надійних методів аналізу, ці сигнали можуть бути використані для оцінки частоти дихання (ЧД, англ. respiratory rate (RR)). Така функціональність значно розширює кількість фізіологічних змінних, які можуть вимірюватися за допомогою споживчих пристроїв, відкриваючи нові можливості для їх застосування.

Крім того, низка останніх досліджень описує підходи до вимірювання частоти серцевих скорочень і частоти дихання без використання сенсорів чи електродів, прикріплених до пацієнта. Замість цього використовуються стандартні відеодані для аналізу фізіологічних параметрів [21, 22, 23]. Одним із результатів таких підходів є хвиля "ВФПГ" (відео-генерована фотоплетизмографія), яка дозволяє оцінювати частоту дихання. Ці методи зазвичай базуються на відстеженні невеликих ділянок відкритої шкіри у відеоматеріалі суб'єкта. Однак сигнал ВФПГ

має значно більше шуму порівняно з традиційним сигналом ФПГ, отриманим від пульсоксиметрів, що створює додаткові технічні виклики для точного аналізу.

У своїй роботі я пропоную альтернативний підхід, побудований на носійних сенсорах (ФПГ), адаптивному фільтрі Калмана та моделях LSTM/1D-CNN, який дозволяє у режимі прикінцевих обчислень попередньо обробити та класифікувати сигнали без необхідності багатоканальної полісомнографії. Така система забезпечує комфорт пацієнта, автономність моніторингу та можливість масового скринінгу поза межами спеціалізованих лабораторій.

2.4.2. Визначення джерел даних

Серед сучасних методів вимірювання частоти дихання та діагностики апное виділяється кілька підходів, кожен із яких має свої переваги, обмеження та специфічні сфери застосування. Одним із ефективних методів є інфрачервона термографія, яка використовує інфрачервоні камери для реєстрації змін температури повітря навколо носа та рота під час вдиху і видиху [64]. Цей підхід дозволяє отримати детальну картину температурних коливань, де підвищення температури відповідає фазі видиху, а зниження – фазі вдиху. Попри свою ефективність, метод має низку обмежень. Для точного застосування необхідно забезпечити стабільні умови навколишнього середовища, оскільки зовнішні фактори, такі як температурні коливання або протяги, можуть знижувати точність вимірювань. Крім того, висока вартість інфрачервоного обладнання та обмеження його застосування у стаціонарних умовах ускладнюють інтеграцію цього методу у мобільні пристрої.

Альтернативним підходом до моніторингу дихання є капнографія, яка аналізує концентрацію вуглекислого газу (CO_2) у повітрі, що видихається. Цей метод є високоефективним у клінічних умовах, зокрема під час реанімаційних заходів або анестезії, оскільки забезпечує точні дані про частоту дихання та якість газообміну [65]. Дані, отримані через капнографію, представляються у вигляді графіка концентрації CO_2 , де кожен пік відповідає фазі видиху пацієнта. Проте капнографія має низку обмежень. Вона потребує спеціалізованого обладнання, яке зазвичай використовується з носовими канюлями або під час інтубації. Це робить

метод інвазивним, а також пов'язаним із додатковими витратами на витратні матеріали. Іншим методом моніторингу дихання є акустичний аналіз, який базується на реєстрації звуків дихання за допомогою мікрофонів. Цей підхід дозволяє отримати інформацію про ритм дихання, але точність результатів значною мірою залежить від правильного розташування мікрофона та умов навколишнього середовища [66]. Навколишній шум може суттєво спотворювати результати. Незважаючи на неінвазивність і простоту акустичного аналізу, цей метод є менш придатним для діагностики апное, особливо у випадках, коли пацієнт дихає майже беззвучно під час сну.

Метод рухових датчиків, таких як акселерометри та гіроскопи, застосовується для вимірювання частоти дихання завдяки здатності відстежувати рухи грудної клітки та живота, спричинені диханням. Ці датчики можуть бути інтегровані в носимі пристрої, що забезпечує зручність використання в умовах повсякденного життя. Проте точність цього методу знижується під час надмірних рухів тіла, які не пов'язані з диханням, що обмежує його ефективність у динамічних умовах. Ще одним непрямим, але перспективним методом діагностики апное є електроенцефалографія (ЕЕГ), яка вимірює біопотенціали мозку, що змінюються під час епізодів апное. Аналіз ЕЕГ дозволяє виявити порушення активності мозку, спричинені апное. Однак цей підхід не є прямим методом вимірювання дихання і краще підходить для виявлення серйозних порушень дихальної функції. Крім того, обладнання для ЕЕГ є складним і дорогим, а точний аналіз потребує кваліфікованого персоналу. Для прямого вимірювання потоку повітря широко використовуються носові канюлі. Цей метод забезпечує високу точність даних про частоту дихання та є особливо корисним для діагностики апное. Канюлі вставляються в ніс пацієнта, реєструючи потік вдихуваного та видихуваного повітря. Попри високу ефективність, цей підхід є інвазивним, викликає дискомфорт і більше підходить для короткочасного моніторингу, ніж для тривалого використання.

Ще одним підходом до реєстрації дихання є використання п'єзоелектричних сенсорів, які розміщуються під матрацом для моніторингу дихання під час сну. Ці

сенсори реєструють вагові коливання, спричинені диханням і серцебиттям, без необхідності прямого контакту з пацієнтом. Попри неінвазивність і зручність використання, точність сигналу може знижуватися у разі змін положення пацієнта або у випадках значної ваги тіла, що спотворює результати вимірювань. Таким чином, кожен із розглянутих методів моніторингу дихання має свої переваги та обмеження. Вибір конкретного підходу залежить від умов застосування та потреб моніторингу, враховуючи точність, зручність, інвазивність і тривалість використання.

Ключовим аспектом визначення подій апное за допомогою ФПГ є аналіз змін амплітуди пульсової хвилі (АПХ) та супутніх десатурацій SpO_2 . Під час пауз дихання обструктивного апное спостерігається виразне зниження АПХ через короткочасне припинення вентиляції та зменшення периферійного кровотоку, тоді як крива насичення киснем демонструє різке падіння нижче критичного рівня (зазвичай 90%). На графіку (рис. 3) показано ці закономірності: на верхньому графіку – криву SpO_2 із двома інтервалами десатурації до 87%, а на нижньому – відповідний PPG-сигнал із пригніченими хвилями в тих самих часових проміжках. Поєднання зменшення АПХ та падіння SpO_2 слугує надійним індикатором подій апное і використовується для автоматизованого розрахунку індексу апное-гіпопное (АНІ), що дозволяє класифікувати тяжкість стану пацієнта.

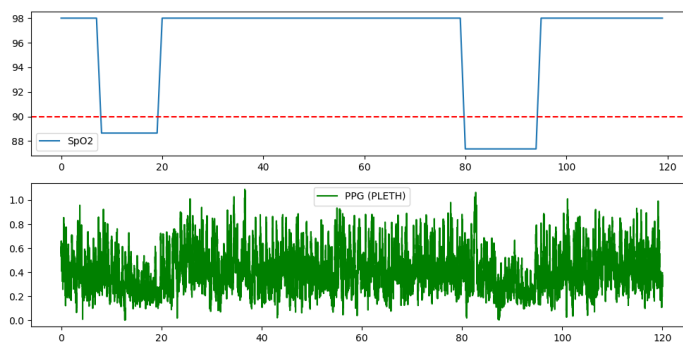


Рисунок 2.1 Десатурації SpO_2 та пригнічена амплітуда PPG під час важкого обструктивного апное сну.

Впровадження алгоритмів аналізу ФПГ-сигналів для виявлення апное сну передбачає суворе дотримання етичних та соціальних норм, починаючи із захисту медичних даних: всі записи ФПГ і результати обробки мають передаватися та зберігатися з використанням сучасних протоколів шифрування (TLS, AES-256) відповідно до вимог GDPR та HIPAA, а пацієнт завчасно підписує інформовану згоду, у якій визначено обсяг і тривалість зберігання даних та порядок їх видалення. Одночасно алгоритми класифікації повинні бути прозорими для лікарів і пацієнтів, із можливістю відстежити, як саме обчислюється індекс АНІ на основі змін амплітуди PPG та падінь SpO₂, та як обрані порогові значення впливають на чутливість і специфічність системи в різних демографічних групах. Вкрай важливо також мінімізувати алгоритмічні упередження: модель тестується на пацієнтах із різним фототипом шкіри і судинною мікроциркуляцією, а будь-які розбіжності в показниках (чутливості чи специфічності) понад 5 % між групами приводяться до відповідної корекції. Водночас усі автоматичні рішення мають бути підзвітними та логуватися з часовими мітками та ідентифікаторами користувачів, що забезпечує аудит і відповідальність у разі сумнівних результатів. Нарешті, хоча система здатна видавати попередню оцінку подій апное, остаточне тлумачення та діагноз належать лікарю, а пацієнт має право відкликати згоду та отримати повний архів своїх даних у зручному для передачі форматі — це зберігає автономію пацієнта та зміцнює його довіру до нових технологій у медицині.

Розміщення громіздких або адгезивних датчиків, таких як електроди ЕКГ, є незручним для тривалого моніторингу через дискомфорт і обмеження мобільності. У зв'язку з цим альтернативним підходом до визначення частоти дихання стає прогнозування цього показника на основі даних про серцевий ритм, пульс, рівень насичення киснем у крові (SpO₂) та параметри фотоплетизмографії. Ці параметри мають високу кореляцію з частотою дихання, що робить їх корисними для непрямого прогнозування. Збір необхідних даних можливий за допомогою мініатюрних сенсорів, які можна інтегрувати в носимі пристрої, наприклад, розмістити на зап'ясті або в одязі. Одним із прикладів є датчик MAX30102, який широко використовується в комерційних пристроях завдяки своїм невеликим

розмірам і енергоефективності. Однак точність його показників може знижуватися за умов яскравого зовнішнього освітлення або під час активних рухів. На основі зібраних даних, використовуючи рекурентні нейронні мережі (наприклад, LSTM) або згорткові нейронні мережі (CNN), можна аналізувати часові ряди для точного прогнозування частоти дихання. Цей підхід дозволяє створювати ефективні системи моніторингу, які є зручними для повсякденного використання і можуть забезпечити високу точність прогнозування.

Пульсоксиметрія є широко визнаним методом для надійного визначення частоти серцевих скорочень та рівня насичення киснем у крові (SpO_2), які базуються на аналізі форми хвилі фотоплетизмографії. Однак точні та стандартизовані методи для оцінки частоти дихання на основі сигналу ФПГ наразі відсутні.

Основною перешкодою для розробки надійних методів оцінки частоти дихання або ЧД у контексті мобільного здоров'я (m-health) є наявність артефактів руху, які суттєво спотворюють сигнал ФПГ. Ці артефакти, спричинені фізичною активністю або зовнішніми умовами, ускладнюють виділення дихальних компонентів із загального сигналу. Крім того, амплітуда та частота дихальних сигналів є значно слабшими порівняно з компонентами, що відповідають серцевій активності, що робить їхній аналіз складнішим. Додаткові виклики включають необхідність розробки алгоритмів, здатних адаптуватися до змін фізіологічних станів пацієнтів, таких як варіації у частоті дихання під час фізичної активності, стресу чи сну. Розробка таких алгоритмів вимагає інтеграції методів фільтрації сигналів, корекції артефактів та застосування машинного навчання для виділення релевантних характеристик PPG-сигналу, пов'язаних із диханням.

Таким чином, створення ефективних методів для оцінки RR із сигналу PPG у сфері m-health залишається відкритим завданням, яке потребує подальших досліджень і міждисциплінарної співпраці. Вирішення цих проблем може значно розширити функціональні можливості мобільних пристроїв для моніторингу здоров'я, забезпечуючи пацієнтам більш повний спектр фізіологічного контролю.

2.4.3. Попередня обробка сигналів фотоплетизмограми.

Фотоплетизмографія є ефективним оптичним методом для неінвазивного моніторингу параметрів серцево-судинної системи, зокрема частоти серцевих скорочень (HR) та рівня насичення киснем у крові (SpO_2). Метод базується на використанні інфрачервоного (ІЧ) та червоного світла, які дозволяють реєструвати зміни об'єму крові в мікро судинах, що виникають внаслідок серцевих скорочень. Однак сигнал ФПГ є чутливим до впливу зовнішніх факторів, таких як рухи тіла, зміни освітлення та артефакти, які можуть значно знижувати якість отриманих даних. У зв'язку з цим, для забезпечення точності вимірювань обов'язковим є застосування процедур попередньої обробки сигналу. Ці процедури включають очищення сигналу від шуму, дрейфу базової лінії та інших спотворень, що досягається за допомогою сучасних методів цифрової фільтрації, таких як смугові фільтри, адаптивна фільтрація або методи на основі аналізу в частотній області.

Фільтрація PPG сигналу необхідна для вирішення кількох важливих завдань:

1. Видалення шуму:

ФПГ сигнал зазнає спотворень під впливом фізіологічних та зовнішніх факторів. До фізіологічних шумів належать рухи тіла, дихання та м'язові скорочення, які додають до сигналу складні компоненти, що не пов'язані з кровотоком. Ці артефакти суттєво спотворюють обчислення частоти серцевих скорочень та насичення крові киснем (SpO_2). Додатково, електромагнітні поля, генеровані іншими пристроями, можуть спричиняти електричний шум, що знижує чистоту PPG сигналу.

2. Дрейф базової лінії:

Зміщення базового рівня сигналу (дрейф) часто виникає через варіації в умовах вимірювання, зокрема зміну температури шкіри, тиск сенсора або його контакт із поверхнею тіла. Дрейф базової лінії може поступово знижувати точність обчислень фізіологічних параметрів. Застосування фільтрів, таких як високочастотні або смугові фільтри, дозволяє компенсувати ці зміщення, забезпечуючи стабільність та надійність сигналу.

3. Розділення корисного сигналу:

Інфрачервоні (IR) та червоні (Red) компоненти сигналу PPG містять ключову інформацію про зміни об'єму крові, що відображають фізіологічні параметри серцево-судинної системи. Ефективна фільтрація сигналу дозволяє виділити корисну інформацію, мінімізуючи вплив шумів, які не мають інформаційного навантаження. Таким чином, зменшення артефактів підвищує точність аналізу та забезпечує достовірність отриманих фізіологічних показників.

Для підвищення точності вимірювань PPG сигналу було застосовано кілька типів фільтрів, спрямованих на обробку сигналів та мінімізацію їхніх спотворень. Нижче наведено ключові методи фільтрації:

- Низькочастотний фільтр використовується для видалення високочастотного шуму, що може виникати через електромагнітні перешкоди або інші зовнішні джерела. Цей тип фільтра дозволяє зберегти стабільність сигналу, видаляючи флуктуації, які не пов'язані з фізіологічними параметрами. Застосування низькочастотного фільтра сприяє забезпеченню чистоти сигналу та збереженню ритмічних коливань, що відображають кровотік.
- Високочастотний фільтр призначений для усунення низькочастотних компонентів, таких як тренди або шуми низької частоти, які спричиняють дрейф базової лінії сигналу. Завдяки цьому фільтру знижуються поступові коливання, залишаючи фізіологічно релевантні компоненти, які відповідають реальним змінам об'єму крові.
- Фільтр ковзного середнього є простим і ефективним методом згладжування сигналу, що базується на обчисленні середнього значення кількох відліків сигналу у ковзному вікні. Цей підхід дозволяє зменшити вплив різких змін у сигналі, забезпечуючи стабільність сигналу при збереженні основних фізіологічних компонентів для подальшого аналізу.
- Фільтр Калмана застосовується для більш складного очищення сигналу, особливо в умовах високого рівня шуму. Цей адаптивний метод забезпечує оптимальну оцінку сигналу, дозволяючи динамічно згладжувати PPG сигнал і зберігати важливі фізіологічні коливання. Завдяки здатності мінімізувати

випадкові спотворення, фільтр Калмана є ефективним інструментом для аналізу ФПГ сигналів у складних умовах вимірювання.

Вибір методів фільтрації ФПГ сигналів залежить від вимог до точності вимірювань, характеристик сигналу та умов збору даних. Часто використовується комбінований підхід, що дозволяє послідовно видаляти різні типи шумів і аномалій, забезпечуючи високу якість отриманих показників. Наприклад, процес обробки сигналу може розпочинатися з використання високочастотного фільтра, який усуває низькочастотні компоненти, такі як повільні коливання або дрейф базової лінії, які не мають відношення до серцево-судинної активності. Наступним етапом є застосування середньозваженого фільтра, який згладжує сигнал, зменшуючи вплив різких змін або аномалій. Завершальним кроком у такій послідовності може бути використання фільтра Калмана. Цей адаптивний алгоритм забезпечує точну реконструкцію сигналу, дозволяючи відновити корисні компоненти, мінімізуючи вплив випадкових шумів і спотворень. Комбінація зазначених методів забезпечує стабільність і достовірність отриманих значень, що є критично важливим для подальшого аналізу фізіологічних параметрів.

Кожен із фільтрів відіграє ключову роль у процесі обробки PPG сигналів, забезпечуючи отримання чистих і достовірних даних для розрахунку основних фізіологічних параметрів. Фільтрація інфрачервоного та червоного сигналів дозволяє ефективно знижувати рівень шуму, усувати дрейф базової лінії та мінімізувати вплив артефактів, викликаних зовнішніми і фізіологічними факторами. Такі підходи до обробки сигналів значно підвищують точність визначення частоти серцевих скорочень та рівня насичення киснем у крові (SpO_2). Завдяки використанню фільтрів можна отримати стабільні показники, які відображають реальні фізіологічні процеси, що робить їх критично важливими для подальшого аналізу та використання в медичних і мобільних системах моніторингу.

Після завершення фільтрації сигнал стає чистим і чітким, зберігаючи періодичні пульсові хвилі, які несуть цінну інформацію про стан серцево-судинної системи. Ці хвилі дозволяють точно обчислити частоту серцевих скорочень — один

із ключових показників, що відображає загальний стан серця і судин. Аналізуючи регулярність та періодичність цих хвиль, система може забезпечити надійний контроль важливих фізіологічних параметрів, що є незамінним у сучасних системах моніторингу здоров'я.

Рівень насичення крові киснем (SpO_2) обчислюється за співвідношенням змінних (AC) та постійних (DC) компонентів фотоплетизмографічних сигналів у двох довжинах хвиль — червоній і інфрачервоній. Якщо позначити амплітуду змінної складової червоного сигналу як $AC_{(Red)}$, його постійну складову — $DC_{(Red)}$, а аналогічні величини в інфрачервоному каналі як $AC_{(IR)}$ і $DC_{(IR)}$, то обчислювальна формула має вигляд:

$$R = \frac{\frac{AC_{Red}}{DC_{Red}}}{\frac{AC_{IR}}{DC_{IR}}} \Rightarrow SpO_2 = A - B \times R$$

де константи А та В підбираються калібруванням (наприклад, $A \approx 104$, $B \approx 17$) відповідно до характеристики сенсора. На рисунку (4) наведено приклад однохвилинного фрагмента ФПГ-сигналів: у верхній частині — Red-канал з позначеними DC-лінім значенням базової лінії та AC-амплітудою, у нижній — IR-канал із відповідними компонентами. Виконана під час обробки фільтрація низьких частот виокремлює DC-складову, а високопрохідний фільтр — AC-складову, що мінімізує вплив рухових артефактів і мережевих наведень. Отримані AC/DC-співвідношення обох каналів у формулі “Ratio of Ratios” забезпечують точний розрахунок SpO_2 та високу надійність даних у медичних системах моніторингу сну.

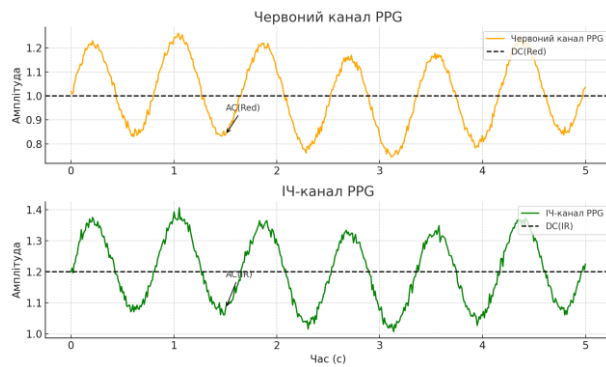


Рис. 2.2. PPG-сигнал: у верхній панелі — червоний канал із базовою лінією DC(Red) і стрілкою, що вказує AC(Red); у нижній панелі — ІЧ-канал із базовою лінією DC(IR) і позначенням AC(IR).

Отже, фільтрація PPG сигналів із використанням комбінації високочастотного фільтра, середньозваженого фільтра та фільтра Калмана є ключовим етапом для підвищення якості отриманих даних. Такий підхід дозволяє значно підвищити точність і надійність розрахунків частоти серцевих скорочень (HR) і рівня насичення крові киснем (SpO_2). Це особливо важливо для ефективного моніторингу фізіологічного стану пацієнтів та прогнозування можливих змін у функціонуванні дихальної системи.



Рис. 2.3 Сирий ІЧ сигнал, записаний без фільтрації, червоні X — піки пульсової хвилі (ЧСС).



Рис. 2.4 Фільтрація ІЧ сигналу низькочастотним фільтром. Синя лінія — сирий ІЧ сигнал, помаранчева — результат низькочастотного фільтра, червоні X — піки ЧСС.



Рис. 2.5 Фільтрація ІЧ сигналу високочастотним фільтром. Синя лінія — сирий ІЧ сигнал, помаранчева — результат високочастотного фільтра, червоні X — піки ЧСС.



Рис. 2.6 Фільтрація ІЧ сигналу середньозваженим фільтром. Синя лінія — сирий ІЧ сигнал, помаранчева — згладжений сигнал середньозваженим фільтром, червоні X — піки ЧСС.



Рис. 2.7 Фільтрація ІЧ сигналу Kalman-фільтром Калмана. Синя лінія — сирий ІЧ сигнал, помаранчева — результат Kalman-фільтрації, червоні хрестики — піки ЧСС.

У наведених на рисунках 2.2.–2.7. ілюстраціях показано, як різні методи фільтрації впливають на первинний ФПГ-сигнал у часі. По осі X час у секундах, по осі Y — амплітуда сигналу в умовних одиницях, що відображає зміну інтенсивності відбитого червоного або інфрачервоного світла під час пульсової хвилі та фаз дихання. Червоні «x» позначають моменти максимального підйому ФПГ-хвилі, які відповідають індивідуальним серцевим скороченням. Кожен із застосованих фільтрів продемонстрував різну ефективність у зменшенні шуму, точності визначення піків та згладженні сигналу, що дозволило оцінити їхні переваги та обмеження.

- Низькочастотний фільтр – усуває повільний дрейф базової лінії та підвищує контраст пульсових піків, що спрощує їх автоматичне виявлення (піки відповідають моментам максимального підйому ФПГ-хвилі — ЧСС). При цьому надмірне згладжування може зменшувати амплітуду короткочасних змін, пов'язаних із десатураціями або початком/завершенням апное, і знижувати чутливість до цих критичних подій.
- Високочастотний фільтр - усуває низькочастотні складові сигналу — дрейф базової лінії та довготривалі тренди, залишаючи лише швидкі коливання, зокрема пульсові хвилі та артефакти. Підсилюються високочастотні шуми та рухові артефакти, що ускладнює їх точне виявлення. Через це фільтр

виявляється менш придатним для надійного розрахунку ЧСС та SpO_2 і детекції подій апное, оскільки короточасні шумові сплески можуть імітувати серцеві піки або приховувати справжні десатурації.

- Середньозважений фільтр – фільтр згладжує високочастотні коливання і усуває випадкові сплески шуму, залишаючи чітко виражені пульсові піки для детекції ЧСС. Проте через згладжування можуть бути зменшені амплітуди коротких десатураційних та апное-епізодів, тому цей фільтр варто використовувати в поєднанні з методами, чутливими до швидких змін.
- Фільтр Калмана - забезпечує адаптивне згладжування, поєднуючи стійкість до шуму з реагуванням на швидкі зміни ФПГ-сигналу. Після фільтрації зберігається чітка форма пульсових піків із мінімальним затриманням, а базова лінія вирівнюється без надмірного згладжування. Це робить фільтр оптимальним для детекції подій апное, оскільки він придушує випадкові шуми, але одночасно не «затирає» короточасні десатураційні епізоди.

Фільтрація PPG-сигналів є обов'язковим етапом підготовки даних для виявлення апное сну: за допомогою низькочастотного фільтра усувається дрейф базової лінії, який маскує справжні респіраторні та пульсові коливання, а високочастотний фільтр видаляє раптові сплески шуму та артефакти руху. Фільтр Калмана, у свою чергу, адаптивно поєднує ці властивості, зберігаючи форму пульсових піків і короточасні десатураційні епізоди навіть у присутності значного рівня перешкод. Такий підхід дозволяє отримати чисті часово-амплітудні характеристики ФПГ — амплітуду пульсової хвилі, частоту серцевих скорочень і респіраторну модуляцію — що є вхідними діагностичними ознаками для автоматичного визначення пауз дихання та обчислення індексу апное-гіпопное (АHI).

Проведений порівняльний аналіз часово-амплітудних характеристик ФПГ-сигналів показав, що з усіх апробованих алгоритмів фільтрації найбільшу ефективність у контексті виявлення пульсових піків і придушення шумових складових демонструють фільтр Калмана та середньозважений фільтр. Зокрема, фільтр Калмана забезпечує адаптивне коригування коефіцієнта посилення шуму,

що дозволяє йому підтримувати високу чутливість до короточасних десатураційних епізодів та швидких коливань амплітуди, водночас знижуючи вплив низькочастотного дрейфу й випадкових сплесків. Середньозважений фільтр, у свою чергу, відзначається відмінною здатністю до усунення високочастотних шумів при збереженні форми основних пульсових хвиль без суттєвих затримок фази.

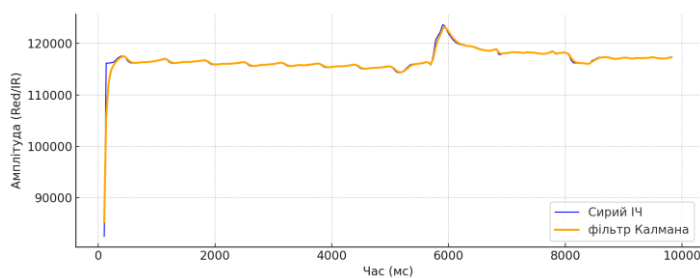


Рис. 2.8. – Порівняння вхідного IR сигналу (блакитний) та відфільтрованих фільтром Калмана (помаранчевий) в спокійному положенні тіла

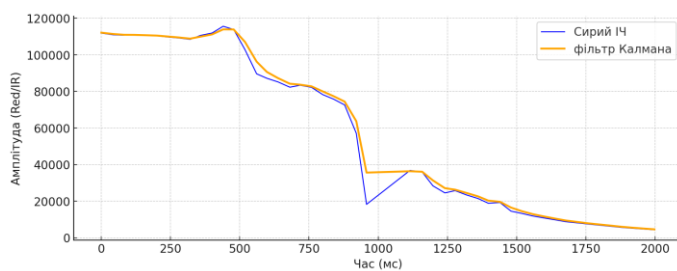


Рис. 2.9. – Порівняння вхідного IR сигналу (блакитний) та відфільтрованих фільтром Калмана (помаранчевий) під час руху

Для виявлення піків у ФПГ-сигналах застосовуються різні автоматичні методи обробки: адаптивне порогове значення [25], методи на основі перетворень [26], обчислення похідних [27] і комп'ютерна фільтрація [28]. Серед цих підходів адаптивні порогові методи є найбільш поширеними для аналізу біомедичних сигналів, таких як ФПГ та ЕКГ. Ці методи встановлюють порогове значення, яке автоматично змінюється залежно від амплітуд виявлених піків. Параметри

порогового значення можуть адаптивно оновлюватися, враховуючи такі характеристики, як тривалість сигналу, амплітуда, інтервали між серцевими ударами та частота вибірки [29, 30]. Водночас, у деяких випадках методи на основі порогових значень вимагають додаткового використання ЕКГ сигналів поряд із ФПГ, що може підвищувати вартість медичного обладнання та ускладнювати його інтеграцію в системи дистанційного моніторингу здоров'я [31]. Попри це, адаптивні техніки залишаються ключовим інструментом для підвищення точності автоматизованого аналізу біомедичних сигналів.

Адаптивний поріг є одним із найпоширеніших методів для виявлення піків у PPG сигналах. Цей підхід базується на використанні змінного порогового значення, яке визначається на основі тимчасових і частотних характеристик сигналу та інтервалів часу. Завдяки динамічному характеру хвильової форми ФПГ, порогове значення може автоматично знижуватися або підвищуватися залежно від умов [25, 32]. Наприклад, Shin та ін. [33] запропонували оновлювати порогове значення залежно від таких параметрів, як частота вибірки, амплітуда попередніх піків і стандартне відхилення сигналу. У дослідженні [34] адаптивний поріг був доповнений морфологічним фільтром для видалення низькочастотного шуму, а функція суми нахилів використовувалася для точного визначення положення піків. Van Gent [35, 36] запропонував метод, який поєднує адаптивний поріг із ковзним середнім і сплайн-інтерполяцією для корекції обрізаних піків. У дослідженні [37] було представлено підхід на основі похідної, оснащений функцією суми нахилів та адаптивним порогом, що спрямований на зменшення кількості помилкових виявлень піків. Методи адаптивного порогу мають низьку обчислювальну складність, що робить їх придатними для використання в реальному часі. Однак їх ефективність значно знижується у випадках, коли сигнал PPG забруднений шумом. У таких ситуаціях швидкі зміни в сигналі ускладнюють точне встановлення порогового значення, що може призводити до помилкових виявлень піків.

На основі виконаних етапів збору та підготовки даних для моделі штучного інтелекту був розроблений комплексний підхід, що забезпечує високу точність і надійність обробленої інформації. Для навчання та тестування моделі було

використано датасет BIDMC PPG and Respiration Dataset [38], доступний у публічному просторі. Цей датасет містить фізіологічні сигнали та числові показники, отримані на основі узгоджених даних із великої бази MIMIC II. Крім того, до нього включені анотації дихання, створені вручну двома незалежними експертами з використанням імпедансного дихального сигналу. Такий підхід до анотації даних гарантує високу якість маркування та надає надійний фундамент для розробки і перевірки моделей штучного інтелекту.

Датасет включає записи про 53 пацієнтів і структурований у три типи файлів: сигнали, числові параметри та загальні дані про пацієнтів, представлені у вигляді фіксованих параметрів. Для забезпечення зручності роботи з даними був розроблений спеціальний скрипт, що об'єднав інформацію про кожного пацієнта в єдиний формат. У остаточному датасеті, підготовленому для аналізу, використовувалися такі ключові параметри: вік, стать, частота серцевих скорочень, рівень насичення крові киснем (SpO_2), респіраторний показник дихання та цифрові сигнали фотоплетизмографії. ФПГ-сигнали зчитувалися з модуля MAX30102 із частотою дискретизації близько 24 Гц (інтервал між відліками ≈ 41 мс) протягом загальної сесії розміром 110 с. Для побудови фізіологічних ознак використовували два відносних вікна: 0–10 000 мс та 0–2 000 мс, достатні для оцінки середньої та миттєвої частоти серцевих скорочень, варіабельності інтервалів частоти дихання, амплітуди пульсових піків для розрахунку SpO_2 і детекції десатурацій під час апное. Така структурована організація даних оптимізує їхнє використання в моделі штучного інтелекту, забезпечуючи повноту й точність подальшого аналізу.

Усі виконані підготовчі етапи, включно з попередньою обробкою сигналів, нормалізацією, обробкою пропущених значень та уніфікацією даних, забезпечують високий рівень якості й узгодженості інформації, необхідної для використання в моделі штучного інтелекту.

РОЗДІЛ 3. РОЗРОБКА МІКРО СЕРВІСНОЇ АРХІТЕКТУРИ СИСТЕМИ МОНІТОРИНГУ І ПРОГНОЗУВАННЯ ДИХАЛЬНИХ РОЗЛАДІВ ПІД ЧАС СНУ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ

Поява Інтернету речей (IoT) відкриває широкі можливості у сфері охорони здоров'я, забезпечуючи інноваційні підходи до моніторингу стану пацієнтів [39, 40]. Зокрема, у контексті телемоніторингу IoT відіграє вирішальну роль як у медичних закладах, так і в особистих пристроях пацієнтів, сприяючи їх інтеграції у повсякденне життя [41, 42]. Розвиток IoT у сфері моніторингу здоров'я має потенціал для значного покращення якості медичних послуг. Це досягається шляхом впровадження систем раннього виявлення захворювань, прогнозування станів пацієнтів та запобігання розвитку небезпечних патологій. Одночасно IoT сприяє оптимізації витрат на медичне обслуговування завдяки зменшенню потреби у дорогих госпіталізаціях [42, 43]. Таким чином, Інтернет речей стає важливим інструментом для трансформації традиційних підходів до охорони здоров'я, забезпечуючи як пацієнтам, так і медичним фахівцям доступ до ефективних, персоналізованих та економічно обґрунтованих рішень.

Лікування хронічних захворювань, догляд за людьми похилого віку та моніторинг фізичної активності під час занять фітнесом — це лише деякі з перспективних медичних застосувань, які можуть бути реалізовані завдяки впровадженню Інтернету речей. У багатьох випадках пацієнти, які проходять лікування, змушені залишатися в лікарні протягом усього періоду терапії, що призводить до значних витрат на госпіталізацію. У цьому контексті технології IoT, які забезпечують віддалений моніторинг стану пацієнтів, пропонують ефективне рішення. Збираючи та передаючи дані про стан здоров'я в режимі реального часу, IoT не лише знижує витрати на медичне обслуговування, але й сприяє ранньому виявленню захворювань та їх ефективному лікуванню [42]. В основі таких рішень лежать розумні медичні датчики, обладнання для візуалізації та діагностики, які є

ключовими елементами IoT у медицині. Крім того, очікується, що впровадження IoT у медичних послугах підвищить їхню ефективність, продовжить життя пацієнтів та надасть їм численні переваги. Для медичних працівників IoT може мінімізувати час простою обладнання завдяки можливості віддаленого надання послуг та моніторингу стану пристроїв у режимі реального часу, що забезпечує оптимізацію ресурсів і підвищення якості медичних процедур.

Медичні IoT-системи генерують значний обсяг даних, особливо під час безперервного моніторингу. Для обробки й зберігання таких даних потрібні відповідні ресурси, включно з довгостроковими архівами. Водночас необхідно забезпечити баланс між збереженням сирих і оброблених даних, щоб уникнути надмірного накопичення інформації. Зберігання медичних даних регулюється суворими нормативами, такими як GDPR [<https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>] або HIPAA [<https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html>]. Для дотримання цих стандартів потрібно впроваджувати політики шифрування даних, анонімізації, регулярного резервного копіювання та управління доступом до інформації. Автоматичний розподіл безперервного потоку ФПГ сигналу на окремі часові вікна та клінічно значущі епізоди, такі як повні серцеві цикли, моменти десатурації або епізоди апное, дає змогу відмовитися від архівації всього необробленого сигналу й зосередитися лише на фрагментах, які мають діагностичне значення, наприклад 10-секундних інтервалів навколо кожного виявленого епізоду апное. Це суттєво спрощує пошук і відновлення конкретних подій, а також гарантує, що в хмарі зберігаються лише ключові сегменти даних, тоді як решту необроблених кадрів можна локально очищати або зберігати на пристрої. Крім того, необхідно гарантувати, що медичні фахівці мають доступ до точних і своєчасних даних, які необхідні для ефективного прийняття рішень.

3.1. Мікросервісна архітектура

Архітектура мікро сервісів представляє собою еволюцію підходу SOA (сервісно-орієнтована архітектура), спрямовану на вдосконалення принципів роз'єднання, гнучкості та розширюваності. Основною концепцією мікросервісів є зменшення спільного використання сервісів і ресурсів, що забезпечує їхню автономність та ефективність [44, 45, 46]. Кожен мікросервіс є невеликим програмним модулем, який відповідає за виконання унікального завдання з мінімальним рівнем відповідальності та слабкою залежністю від інших компонентів. Ця архітектура базується на принципі "ділитися якомога менше" [44], у той час як SOA, навпаки, спрямована на максимізацію спільного використання ресурсів. Мікросервіси забезпечують чітко визначені інтерфейси для взаємодії, що дозволяє досягти високого рівня розділення та розширюваності. Завдяки цьому кожен мікросервіс може оновлюватися незалежно від інших компонентів системи, що значно підвищує гнучкість та спрощує розробку й обслуговування. На відміну від SOA, яка використовує централізовані залежності, такі як загальне сховище даних, мікросервіси керують власними залежностями, включаючи базу даних, сховище ключів і значень, пошукові індекси та черги. Це дозволяє розробляти та розгортати кожен мікросервіс автономно, що є важливою перевагою в умовах динамічних систем [44, 45]. Утім, у реальних сценаріях часто виникає необхідність обміну даними між мікросервісами. У запропонованій архітектурі дані, зібрані одним мікросервісом, можуть бути збережені у спільній базі даних, доступ до якої надається іншому мікросервісу для подальшої обробки [46]. Такий підхід забезпечує баланс між автономністю мікросервісів і потребою в інтеграції даних для досягнення спільних цілей.

Мікросервіси класифікуються на дві основні групи: функціональні мікросервіси (ФМ) та інфраструктурні мікросервіси (ІМ). Функціональні мікросервіси реалізують бізнес-функції, пов'язані з конкретним бізнес-доменом, наприклад, оцінка слабкості системи чи прогнозування ризиків. Вони доступні для зовнішніх клієнтів, таких як носимі пристрої, мобільні телефони та веб-додатки, через шлюз API (інтерфейсу прикладного програмування) (рис. 3.1.). На відміну від

них, інфраструктурні мікросервіси виконують нефункціональні завдання, включаючи ведення журналів, аудит та моніторинг.

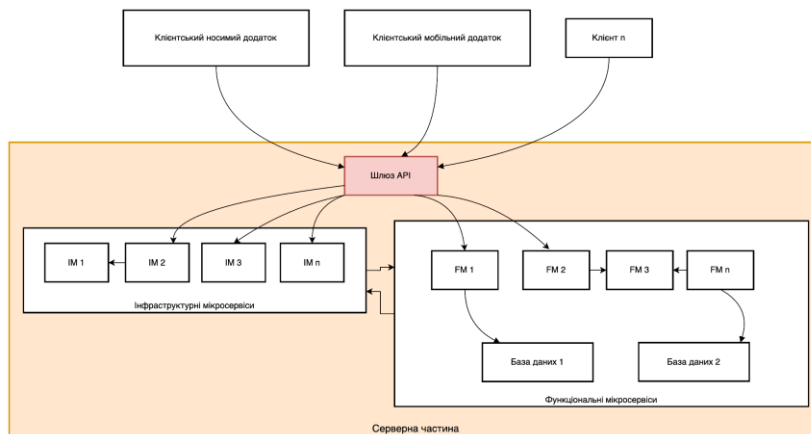


Рис. 3.1. – Таксономія мікросервісної архітектури

Мікросервісна архітектура має низку важливих переваг, які роблять її особливо ефективною для використання в медичних системах:

- **Масштабованість.** Кожен мікросервіс може бути масштабований незалежно, що дозволяє оптимально використовувати ресурси системи. Такий підхід забезпечує стабільно високу продуктивність навіть за умов зростання кількості користувачів або підключених пристроїв, що є критично важливим для медичних систем із великим обсягом даних.
- **Модульність.** Розподіл системи на окремі мікросервіси полегшує інтеграцію нових функціональних можливостей без ризику впливу на інші компоненти. Це особливо важливо для медичних систем, які часто потребують адаптації до нових стандартів, вимог або змін у клінічних протоколах.
- **Гнучкість.** Незалежне оновлення або розгортання мікросервісів дозволяє швидко впроваджувати зміни та вдосконалення без зупинки всієї системи.

Такий підхід мінімізує ризик простоїв і дає можливість оперативно реагувати на зміни у потребах медичних закладів чи пацієнтів.

Завдяки цим перевагам, мікросервісна архітектура стає оптимальним вибором для медичних систем, забезпечуючи стабільність, адаптивність і ефективність у динамічному середовищі сучасної охорони здоров'я.

3.2. Подієво-орієнтована архітектура

Архітектура, керована подіями (англ. Event-Driven Architecture, EDA), є поширеним підходом у сучасних додатках, зокрема в системах, побудованих на основі мікро сервісів. Вона використовує події як засіб запуску процесів і обміну даними між незалежними сервісами. Подія у цьому контексті означає зміну стану або оновлення, наприклад, додавання даних про пацієнта до електронної медичної карти. Події можуть містити деталізовану інформацію про стан (наприклад, назву товару, його ціну та адресу доставки) або слугувати ідентифікаторами змін (наприклад, повідомлення про те, що замовлення було відправлено).

Однією з ключових переваг архітектури є можливість роботи з потенційно нескінченними потоками даних, оскільки запити застосовуються до даних у режимі реального часу. Система обробляє вхідні події негайно, і результати стають доступними відразу після перегляду послідовності подій. Такий підхід забезпечує високоефективну аналітику в реальному часі, що є критично важливим для багатьох сучасних додатків [47].

Подієво-керована архітектура є популярним шаблоном розподіленої асинхронної архітектури, яка широко використовується для розробки високо масштабованих додатків [48]. Завдяки своїй гнучкості цей підхід однаково добре адаптується як для малих систем, так і для великих і складних інженерних проектів.

Архітектура, керована подіями, базується на двох основних топологіях: посередника (mediator) та брокера (broker). Топологія посередника застосовується у випадках, коли необхідно організувати виконання кількох кроків у рамках однієї події. Натомість топологія брокера використовується для організації подій у вигляді ланцюжка, що дозволяє обробляти події без залучення центрального посередника. Це забезпечує більш децентралізований підхід до обробки подій, сприяючи гнучкості та зниженню залежностей між компонентами системи.

У топології медіатора виділяють чотири основні типи архітектурних компонентів (див. Рис. 3.2.): черги подій, медіатор подій, канали подій та процесори подій. Потік подій у цій архітектурі починається з клієнта, який надсилає подію до черги подій. Черга подій слугує механізмом транспортування, передаючи подію до медіатора подій. Медіатор подій приймає початкову подію та організовує її обробку, генеруючи додаткові асинхронні події. Ці події надсилаються до каналів подій, через які вони розподіляються для виконання кожного кроку процесу. Процесори подій, які постійно прослуховують канали подій, отримують події від медіатора та виконують відповідну бізнес-логіку для їхньої обробки. Цей підхід забезпечує гнучкість і модульність системи, дозволяючи легко адаптувати або змінювати окремі компоненти без впливу на всю архітектуру.

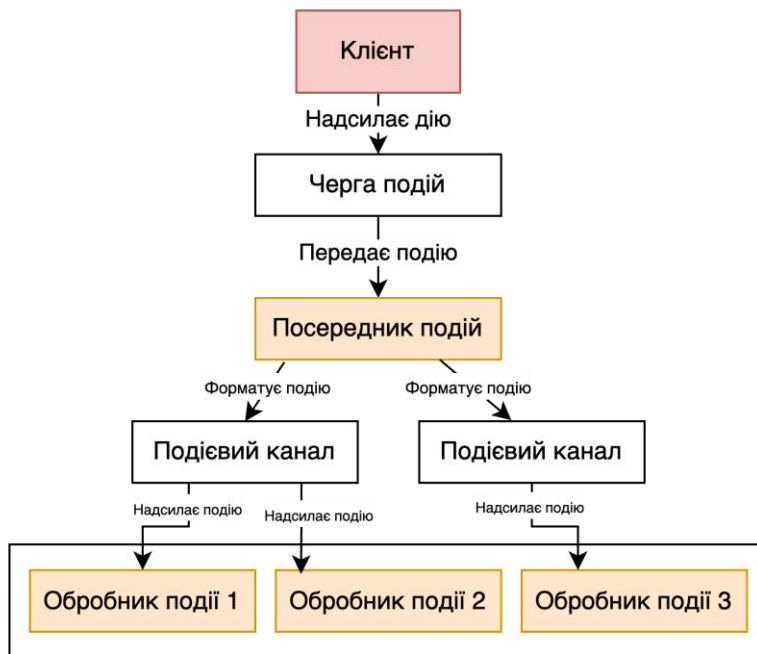


Рис. 3.2. – Топологія медіатора EDA

Топологія брокера суттєво відрізняється від топології посередника тим, що в ній відсутній центральний посередник подій. Замість цього потік повідомлень організовується через ланцюжок компонентів обробки подій за допомогою легкого брокера повідомлень, такого як RabbitMQ [<https://www.rabbitmq.com/>], Kafka [<https://kafka.apache.org/>] тощо. Ця топологія є ефективною для сценаріїв, де потрібно організувати відносно простий потік обробки подій. Вона особливо корисна в системах, де основний акцент робиться на мінімізацію залежностей між компонентами та забезпечення масштабованості.

У брокерській топології виділяють два основних типи компонентів архітектури (див. Рис. 3.3.): компонент брокера та компонент процесора подій. Компонент брокера може бути реалізований як централізована або федеративна структура. Він відповідає за управління всіма каналами подій, які використовуються у потоці подій. Канали подій, що знаходяться в компоненті брокера, можуть бути представлені у вигляді черг повідомлень, тем повідомлень, або їх комбінації, залежно від специфіки обробки та передачі даних у системі. Такий підхід забезпечує гнучкість і модульність у налаштуванні потоків подій, дозволяючи адаптувати систему до різних сценаріїв використання.

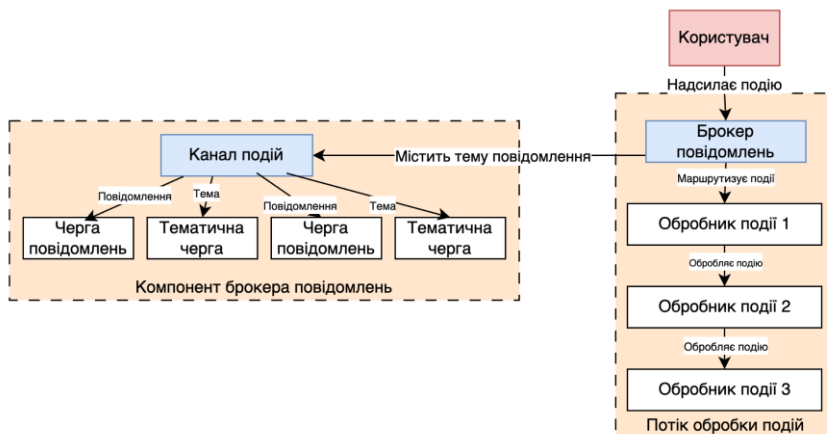


Рис. 3.3. – Топологія брокера EDA

У мікросервісній архітектурі для моніторингу ФПГ-сигналів і прогнозування апное модель «публікація–підписка» виступає єдиною шиною обміну подіями між незалежними сервісами. Така організація забезпечує одночасну доставку одного і того ж пакета даних усім обробникам, що суттєво знижує затримки, підвищує надійність і дає змогу легко масштабувати систему, додаючи нові підписники без змін у коді публікатора. Оскільки кожен вихідний канал обслуговує лише один підписник, ми гарантуємо унікальне споживання повідомлення та строгий контроль за розподілом медичних даних у реальному часі.

Подієво-орієнтована архітектура реального часу може обробляти потоки даних одразу після виникнення подій, забезпечуючи асинхронну взаємодію слабо зв'язаних компонентів і легке горизонтальне масштабування [49]. Саме ці властивості роблять ПОА доцільною основою для медичних інформаційних систем, де затримка усього в кілька секунд може вплинути на клінічний результат. У нашій системі класифікації апное ПОА використано як каркас, у якому мобільний додаток генерує подію «показник обчислено», а мікросервіс штучного інтелекту — «ризик апное класифіковано». Кожен модуль може оновлюватися без зупинки всієї платформи, а відмова одного сервісу не блокує інші, що підвищує надійність і дозволяє поступово додавати нові аналітичні алгоритми під зростаючі обсяги даних. Таким чином, саме завдяки властивим ПОА швидкій реакції, гнучкості й відмовостійкості вдалося реалізувати запропоновану в дисертації систему домашнього моніторингу апное, яка безперервно приймає та обробляє дані від багатьох носимих пристроїв без ризику перевантаження чи втрати критичної інформації.

3.3. Прикінцеві обчислення

Прикінцеві обчислення — це модель інформаційних обчислень, що передбачає перенесення функцій обробки та зберігання даних із централізованих центрів обробки даних на сервери, розташовані ближче до межі мережі, де користувачі та пристрої отримують доступ до інформації. Завдяки розташуванню обчислювальних потужностей ближче до кінцевих користувачів і пристроїв,

периферійні обчислення сприяють значному покращенню продуктивності систем. Цей підхід знижує затримки, забезпечуючи більш швидку передачу та обробку даних, що є критично важливим для додатків у режимі реального часу. Крім того, дана модель дозволяє мінімізувати витрати, оптимізуючи використання ресурсів та мережевих з'єднань, а також посилює безпеку, зменшуючи кількість даних, які передаються до централізованих сховищ. Периферійні обчислення також підвищують якість обслуговування, забезпечуючи швидший доступ до інформації та зменшуючи залежність від інфраструктури центральних обчислювальних центрів [50].

Обсяги даних, які генеруються та використовуються щороку, зростають у геометричній прогресії. Технології, такі як Інтернет речей (IoT), штучний інтелект (ШІ) та машинне навчання (ML), є основними джерелами цього зростання, генеруючи величезні масиви інформації. З підключенням мільярдів пристроїв до мереж 5G, цей обсяг даних продовжуватиме швидко збільшуватись.

У традиційних обчислювальних моделях дані організацій переспрямовуються до централізованих центрів обробки, де ІТ-спеціалісти виконують їхнє управління, перевірку та забезпечують відповідність політикам безпеки. Однак зі зростанням обсягів даних цей підхід стикається з численними обмеженнями. Пересилання всього трафіку до централізованого центру обробки неминуче викликає затримки та перевантаження мереж. Це негативно впливає на продуктивність додатків та швидкість доступу до даних, що призводить до незадовільного користувацького досвіду. Усе це підкреслює необхідність нових підходів, здатних ефективно працювати в умовах експоненційного зростання обсягів інформації.

Edge Computing ефективно вирішує проблему перевантаження традиційних обчислювальних моделей, обробляючи та зберігаючи дані ближче до місця їхнього створення та використання. Розташовуючи обчислювальні функції, робочі навантаження та дані поблизу кінцевих пристроїв, периферійні обчислення дозволяють значно покращити час відгуку, продуктивність додатків і загальний досвід користувачів (див. Рис. 3.4.). Завдяки цьому дані можуть швидше досягати

кінцевих точок, мінімізуючи затримки та забезпечуючи більш ефективну роботу додатків у реальному часі. Паралельно периферійні обчислення зменшують перевантаження центральних систем і ширших мереж, сприяючи стабільності та оптимізації інфраструктури.

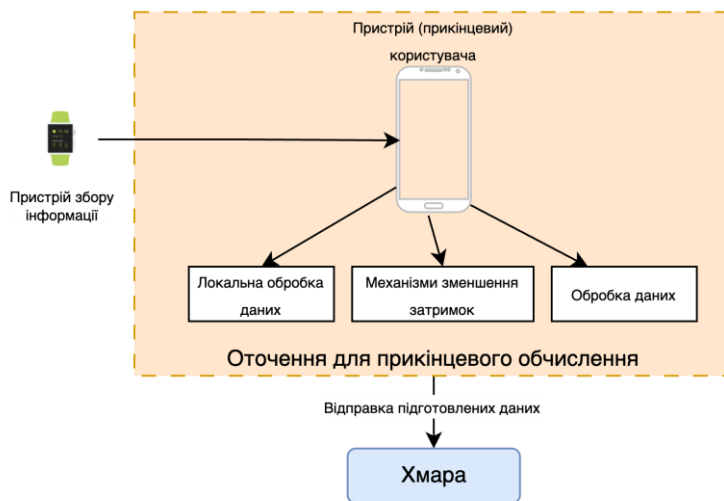


Рис. 3.4. – Принцип роботи моделі прикінцевих обчислень

З розвитком технологій Інтернету речей (IoT) різні галузі активно впроваджують цифрову трансформацію, спрямовану на інтеграцію дедалі більшої кількості пристроїв до глобальної мережі. За прогнозами, до 2025 року кількість підключень IoT-пристроїв у світі досягне 100 мільярдів. Така швидка експансія IoT створює значні можливості для підприємств, але водночас ставить перед ними низку викликів, серед яких:

- Коли величезні обсяги даних переносяться в хмару для обробки, відсутність можливостей аналізу та обробки даних в режимі реального часу значно збільшує навантаження на хмару.
- Важко централізовано розгортати і керувати численними пристроями і додатками Інтернету речей, а також різноманітними інтерфейсами і протоколами.

Для подолання цих викликів підприємствам доводиться розробляти інноваційні рішення, які враховуватимуть складність IoT-середовищ і потребу в адаптації до нових технологічних умов.

Периферійні обчислення розгортаються на межі мережі, поблизу джерел даних або пристроїв, і забезпечують послуги аналізу за допомогою відкритої платформи. Ця платформа інтегрує можливості мереж, обчислювальних ресурсів, сховищ даних і додатків, створюючи єдине рішення для обробки інформації. Дані, які збираються термінальними пристроями, аналізуються та обробляються локально на межі мережі в режимі реального часу. Такий підхід усуває необхідність пересилання великих обсягів даних до хмарних центрів для подальшої обробки, що значно знижує затримки та покращує швидкість реагування системи. Граничні обчислення ідеально відповідають ключовим вимогам цифровізації промисловості, забезпечуючи гнучке підключення пристроїв і оптимізацію даних у режимі реального часу. Вони створюють передумови для ефективної роботи систем, де критично важливими є низькі затримки, висока швидкість обробки даних і мінімальна залежність від централізованих ресурсів [51].

Поєднання периферійних обчислень і технологій IoT створює новий підхід — периферійні обчислення IoT. Цей підхід визначає архітектуру периферійних обчислень у контексті IoT-систем, інтегруючи можливості мереж, обчислень, зберігання даних і додатків. Ключовим елементом цієї архітектури є шлюз периферійних обчислень, який розгортається на межі мережі поблизу пристроїв або джерел даних. Шлюз забезпечує послуги з управління та контролю пристроїв, що працюють на периферійних вузлах мережі. Такий підхід дозволяє вирішити проблему «останньої милі» промислового зв'язку IoT — вузького місця, де часто виникають складнощі з ефективним обміном даними між кінцевими пристроями та основною інфраструктурою. Периферійні обчислення IoT забезпечують інтелектуальне підключення та ефективне управління пристроями IoT, що сприяє підвищенню швидкості, надійності та продуктивності в промислових і споживчих середовищах.

3.4 Комунікація між компонентами

3.4.1 Bluetooth низького енергоспоживання

Bluetooth Low Energy (BLE) розроблений для забезпечення бездротової комунікації на близьких відстанях із мінімальними енергозатратами. Ця технологія дозволяє пристроям, таким як ESP32, працювати протягом тривалого часу без необхідності частої заміни чи заряджання батарей, що є критично важливим для медичних пристроїв, які забезпечують безперервний моніторинг стану пацієнтів. Завдяки високій енергоефективності, ESP32 може постійно збирати та передавати дані про стан здоров'я, не спричиняючи значного виснаження батареї. Це дозволяє використовувати пристрої у тривалих моніторингових процесах без втрати їхньої функціональності. BLE підтримує кілька режимів роботи, включаючи періодичне передавання даних (advertising) та встановлення постійного з'єднання між пристроями. Така гнучкість дозволяє налаштовувати комунікацію відповідно до потреб системи, забезпечуючи баланс між продуктивністю, енергоспоживанням і вимогами до передачі даних [52].

Bluetooth Low Energy (BLE) використовує концепцію GATT (Generic Attribute Profile) для організації передачі даних між пристроєм ESP32 та мобільним телефоном. У цій моделі ESP32 виступає як периферійний пристрій (peripheral), що рекламує свої послуги (services) та характеристики (characteristics), тоді як мобільний телефон виконує роль центрального пристрою (central), який ініціює з'єднання та запити даних. Кожна характеристика може представляти певний параметр здоров'я, наприклад, рівень кисню в крові або частоту серцевих скорочень. Дані можуть бути запитані мобільним додатком періодично або передаватися автоматично, коли ESP32 надсилає оновлення через механізм повідомлень (notifications). Це дозволяє передавати медичні показники у режимі реального часу, забезпечуючи мінімальну затримку.

Крім своєї енергоефективності, BLE забезпечує низьку затримку та можливість передачі даних у реальному часі, що є критичним для медичних систем. Швидка реакція на зміни в стані пацієнта, наприклад, у разі різкого зниження рівня

кисню в крові, дозволяє передати дані через BLE на мобільний пристрій для подальшої обробки або миттєвого повідомлення лікаря. Ще однією перевагою BLE є його гнучкість налаштувань. Технологія може використовуватися для регулярної передачі невеликих обсягів даних, наприклад, базових медичних показників, або для роботи з більшими обсягами даних, такими як повні логи фізіологічних параметрів. Це робить BLE універсальним рішенням для різноманітних задач у медичних системах.

3.4.2. Асинхронна передача повідомлень між мікросервісами.

У реалізованій архітектурі мобільний застосунок спочатку пакує 30с фрагменти ФПГ-даних у повідомлення й публікує їх у чергу через MQTT-брокер, що дозволяє зі зниженим енергоспоживанням і мінімальним мережевим навантаженням передавати сирі цифрові сигнали фотоплетизмографії [53]. Далі ці повідомлення потрапляють до внутрішньої системи обміну подіями на базі RabbitMQ, де кожне повідомлення автоматично копіюється в черги для відповідних мікросервісів. Після чого очищені дані надходять у сервіс прогнозування на основі моделі штучного інтелекту. Коли класифікація завершена, результати — індекс АНІ та клас апное — публікуються в окрему чергу RabbitMQ, звідки мобільний клієнт, підписаний на цю чергу, отримує повідомлення про готовність результатів і відразу відображає їх користувачеві. Така комбінація MQTT для периферійного зв'язку та RabbitMQ для серверної шини гарантує одночасну доставку даних усім необхідним сервісам із низькою затримкою, високою надійністю та можливістю гнучкого масштабування без жодних блокувань між компонентами.

3.4.3. Використання інтерфейсу прикладного програмування для взаємодії між мобільним додатком та ІІІ шлюзу.

HTTP/REST API є одним із ключових механізмів для забезпечення взаємодії між мобільними додатками та серверними компонентами, зокрема через API Gateway. У контексті медичної IoT-системи, яка використовує ESP32 для збору

даних, REST API виконує роль централізованої точки входу для обміну інформацією між мобільними додатками та серверними мікросервісами.

Цей підхід має кілька переваг:

- Гнучкість управління запитами. REST API дозволяє легко адаптувати та керувати потоками даних між різними компонентами системи, зокрема мобільними додатками, IoT-пристроями та мікросервісами.
- Безпека передачі даних. Завдяки використанню стандартів безпеки, таких як HTTPS та аутентифікація через токени, REST API забезпечує конфіденційність та захист чутливих медичних даних.
- Масштабованість. Архітектура REST API дозволяє ефективно обробляти зростаючий обсяг запитів, що надходять від IoT-пристроїв та мобільних додатків, забезпечуючи стабільну роботу системи навіть за значного навантаження.

Використання REST API у медичних IoT-системах забезпечує інтеграцію між різними компонентами інфраструктури, дозволяючи реалізувати масштабовані, надійні та безпечні рішення для обробки медичних даних.

API Gateway функціонує як єдиний вхід для всіх запитів від мобільних додатків і взаємодіє з іншими мікросервісами у хмарній системі. Це надає кілька ключових переваг:

- Централізоване управління запитами: API Gateway приймає всі HTTP/REST-запити від мобільного додатка і маршрутизує їх до відповідних мікросервісів. Це полегшує управління взаємодією між різними компонентами системи, знижує складність мобільного додатка і централізує обробку.
- Безпека: API Gateway може інтегрувати механізми шифрування та автентифікації, забезпечуючи безпечну передачу медичних даних, зокрема через шифрування за допомогою TLS. Це особливо важливо для медичних даних, оскільки конфіденційність і цілісність інформації є критичними вимогами.
- Масштабованість: API Gateway допомагає забезпечити масштабованість системи, оскільки він може розподіляти запити між різними мікросервісами.

Наприклад, коли кількість запитів збільшується, можна додати нові інстанції мікросервісів, а API Gateway автоматично розподілить навантаження між ними.

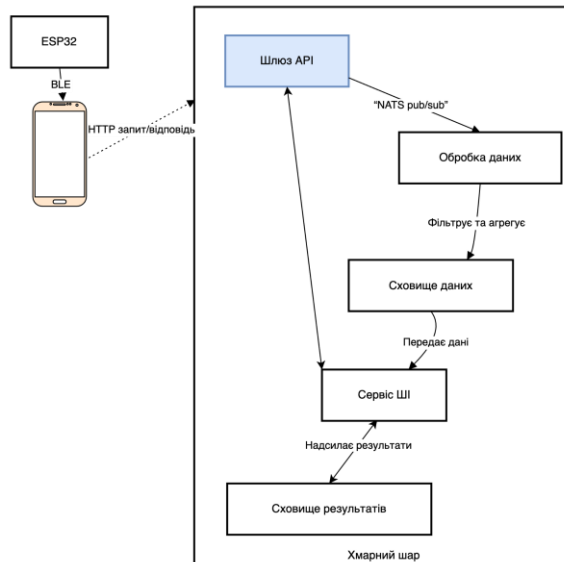


Рис. 3.5. – Архітектура системи збору та аналізу захворювань дихання

На діаграмі (рис. 3.5.) представлено архітектуру медичної IoT-системи, яка ілюструє повний шлях передачі даних від IoT-пристрою ESP32 до хмарних сервісів, де ці дані обробляються та зберігаються. Система починається з ESP32, який збирає медичні показники, наприклад, рівень кисню в крові (SpO_2) або частоту серцевих скорочень (HR), за допомогою підключених сенсорів. Для забезпечення ефективної передачі даних пристрій використовує протокол Bluetooth Low Energy (BLE), що гарантує енергоефективність і стабільність передачі навіть у режимі тривалого моніторингу. Дані надсилаються до мобільного додатка, який виступає як центральна точка для збору інформації, а також для її первинної обробки чи передачі на сервер. Завдяки такій архітектурі забезпечується безперервна взаємодія між пристроєм ESP32 та хмарними сервісами, що дозволяє реалізувати моніторинг

стану пацієнта в режимі реального часу та підвищити якість медичного обслуговування (рис. 3.6).

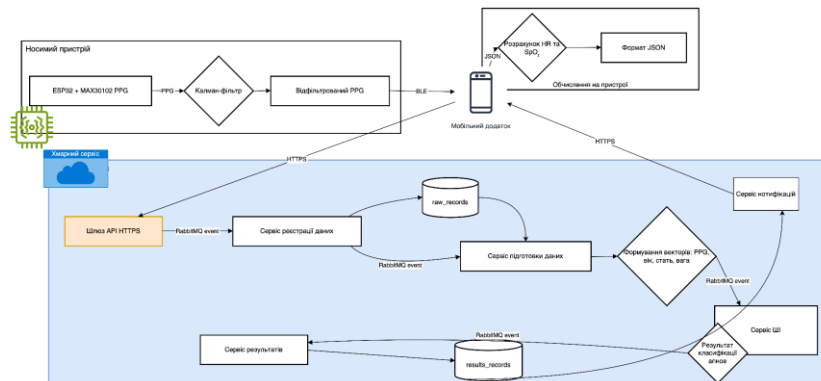


Рисунок 3.6 Схема роботи системи

Мобільний додаток приймає зібрані пристроєм ESP32 медичні дані, які потім трансформуються у формі HTTP-запитів для передачі до хмарної системи через API Gateway. Ця компонента виступає центральною точкою входу до системи, де виконується маршрутизація запитів до відповідних мікросервісів. У процесі обробки даних API Gateway забезпечує автентифікацію та авторизацію користувачів, що гарантує безпеку та захист чутливої інформації. Завдяки такій архітектурі забезпечується надійний і масштабований обмін даними між IoT-пристроями, мобільним додатком та хмарними сервісами. Це є ключовим для медичних IoT-систем, які вимагають високого рівня захисту та ефективності в обробці даних.

Після передачі через API Gateway дані потрапляють до сервісу обробки даних, де відбувається їх очищення та фільтрація для забезпечення якості та достовірності інформації. Цей етап гарантує, що отримані медичні показники готові до подальшого використання в аналізі та прийнятті рішень. Оброблені дані зберігаються в сервісі зберігання показників, який використовує спеціалізовані бази даних для організації надійного зберігання інформації про медичні показники пацієнтів. Паралельно дані передаються до сервісу штучного інтелекту, де вони аналізуються за допомогою моделей машинного навчання. Цей процес дозволяє

виконувати точне прогнозування можливих ризиків для здоров'я пацієнтів, підвищуючи якість медичного моніторингу та забезпечуючи своєчасне виявлення потенційних загроз.

Результати аналізу та прогнозування зберігаються у спеціалізованому сервісі запису результатів, що дозволяє лікарям і пацієнтам отримувати доступ до цих даних для подальшого перегляду та прийняття рішень. Така архітектура забезпечує надійний, масштабований та ефективний потік даних, починаючи з їх збирання та завершуючи їх зберіганням і аналізом у хмарній системі.

Використання BLE (Bluetooth Low Energy) для передачі даних між пристроєм ESP32 і мобільним додатком дозволяє досягти низького рівня енергоспоживання, що є критично важливим для забезпечення тривалої автономної роботи IoT-пристроїв. Завдяки мобільному додатку, який виконує ключову роль у передачі даних до хмарних сервісів через API Gateway, система забезпечує централізоване управління запитами, а також конфіденційність і безпеку переданих медичних показників.

3.5. Розробка архітектури системи моніторингу і прогнозування дихальних розладів під час сну на основі штучного інтелекту

Система побудована на основі мікросервісної архітектури, що забезпечує високу гнучкість і масштабованість, а також використовує подієво-орієнтовану архітектуру (EDA) для асинхронної обробки даних. Додатково впроваджено обчислення на периферії (edge computing), що дозволяє виконувати попередню обробку даних безпосередньо на пристроях або поблизу них. Ця архітектура охоплює всі ключові етапи роботи з інформацією: від збору даних IoT-пристроями, таких як ESP32, до їх обробки, передачі, зберігання у хмарній інфраструктурі та інтеграції з мобільними додатками. Завдяки такій організації забезпечується ефективний та безперервний потік даних, що підтримує високий рівень продуктивності та надійності системи.

3.5.1. Збір даних

На етапі збору даних у системі застосовується концепція прикінцевих обчислень, що дозволяє виконувати початкову обробку інформації безпосередньо на пристрої. IoT-пристрій ESP32, оснащений медичними сенсорами, здійснює моніторинг ключових фізіологічних показників, таких як SpO_2 і серцевий ритм, у режимі реального часу. Оскільки ESP32 має обмежені обчислювальні ресурси, попередня обробка даних виконується прямо на пристрої. Це включає базову фільтрацію сигналів, яка зменшує рівень шуму та покращує якість переданих показників. Такий підхід дозволяє зменшити обсяг даних, які потрібно передавати до хмарної системи, що знижує навантаження на мережу і підвищує швидкість реагування всієї системи. Завдяки попередній обробці даних на пристрої забезпечується баланс між автономністю IoT-пристрою та ефективністю аналізу медичних показників.

3.5.2 Обробка даних на рівні периферії

Після збору медичних даних пристроєм ESP32 інформація передається до мобільного додатка через Bluetooth Low Energy (BLE). Мобільний додаток також використовує концепцію edge computing, забезпечуючи додаткову попередню обробку даних. На цьому етапі додаток може виконувати очищення або агрегування даних, що дозволяє підготувати їх до передачі у хмарну інфраструктуру. Така розподілена обробка даних значно знижує навантаження на хмарну систему, оптимізуючи обробку великих обсягів інформації. Це дає змогу забезпечити стабільну роботу системи навіть за умови одночасного використання багатьма IoT-пристроями. Завдяки попередній обробці на мобільному додатку досягається баланс між локальною ефективністю та швидкістю передачі даних у хмару, що є ключовим для медичних IoT-систем, які працюють у режимі реального часу.

3.5.3 Відправка даних в хмару

Для передачі даних у хмарну систему мобільний додаток використовує HTTP/REST API, який працює через API Gateway. Цей компонент виконує роль

центральної точки входу для всіх запитів із додатка, забезпечуючи безпечну передачу даних до хмарної інфраструктури. API Gateway дозволяє централізовано керувати процесами автентифікації та авторизації, що гарантує захист конфіденційної медичної інформації пацієнтів. У хмарній системі завдяки впровадженню мікросервісної архітектури кожен мікросервіс виконує чітко визначену функцію в процесі обробки даних. Такий підхід забезпечує модульність і масштабованість системи, дозволяючи ефективно адаптувати її до зростаючого навантаження або інтегрувати нові функції без порушення роботи інших компонентів.

3.5.4. Зберігання даних в хмарі

До хмарної інфраструктури передаються відфільтрований сигнал ФПГ із часовими мітками (30-секундний фрагмент при частоті ~ 100 Гц), а також послідовність міток систолічних піків, на основі яких формується beat-to-beat ряд ЧСС із приблизно 30–45 значеннями за 30с, і ковзні оцінки сатурації, розраховані в 5-секундних вікнах із кроком 1с (≈ 26 значень за той самий проміжок). Ці дані зберігаються в спеціалізованих базах даних, які інтегровані в рамках мікросервісної архітектури. Для збереження медичних показників, представлених у вигляді часорядів, використовуються Time-Series бази даних, наприклад, InfluxDB. Такі бази оптимізовані для роботи з великими обсягами безперервних даних, що дозволяє швидко отримувати доступ до історичних показників пацієнтів. Кожен мікросервіс у системі має можливість працювати зі своєю окремою базою даних. Цей підхід забезпечує гнучкість у масштабуванні та дозволяє ефективно розподіляти обробку між різними сервісами. Завдяки такій організації система здатна легко адаптуватися до зростаючих обсягів даних, забезпечуючи стабільну роботу навіть за умов інтенсивного навантаження.

3.5.5. Обробка даних в хмарі

Після надходження в хмару дані обробляються спеціалізованим мікросервісом, основною функцією якого є очищення та нормалізація інформації.

Цей сервіс готує дані для подальшого аналізу, зберігання або передачі до інших компонентів системи. У процесі обробки можуть виконуватися фільтрація для усунення шумів та агрегування, що сприяє оптимізації використання ресурсів для зберігання та аналізу. Завдяки обчислювальним можливостям хмарної інфраструктури, сервіс здатен ефективно обробляти великі обсяги даних, які надходять одночасно від численних IoT-пристроїв. Такий підхід дозволяє забезпечити високу продуктивність системи, зберігаючи якість і цілісність інформації для подальшого використання у прогнозуванні або прийнятті рішень.

3.5.6. Підготовка даних для відправки до сервісу класифікації

Після завершення попередньої обробки дані передаються до спеціалізованого сервісу штучного інтелекту, який виконує саме класифікацію епізодів обструктивного апное сну та визначає їхню тяжкість на основі часово-амплітудних ознак ФПГ, варіабельності ЧСС-інтервалів і змін сатурації. У цьому сервісі використовуються моделі машинного навчання, які забезпечують глибокий аналіз отриманої інформації та виявлення потенційних загроз у стані здоров'я пацієнтів. Сервіс функціонує на основі подієво-орієнтованої архітектури, що дозволяє розглядати кожен новий пакет даних як подію. При надходженні нового пакету автоматично запускається процес аналізу, що забезпечує швидке реагування системи на зміну медичних показників пацієнта. Завдяки цьому система здатна генерувати прогнози в реальному часі, що є критичним для своєчасного прийняття рішень у медичних застосуваннях.

3.5.7. Обробка результатів

Після аналізу даних за допомогою моделей штучного інтелекту система формує результати прогнозування, які інтегруються в межах мікросервісної архітектури. Ці результати можуть включати як аналітичні висновки, так і персоналізовані рекомендації для лікарів або пацієнтів. Оброблені дані автоматично передаються до відповідних сервісів для подальшої обробки, зберігання або відображення користувачам. Такий підхід забезпечує гнучкість і

ефективність системи, дозволяючи адаптувати результати до конкретних потреб користувачів. Водночас централізоване зберігання результатів у спеціалізованих сервісах гарантує доступність і безпеку даних для подальшого використання.

3.5.8. Збереження результатів

Результати обробки зберігаються не як загальні «прогнози ризиків», а у вигляді чітких класифікацій подій дихання та обчисленого індексу апное-гіпноное (АНІ). Кожен 30-секундний фрагмент PPG сигналу класифікується одним із п'яти станів:

- Normal — нормальне дихання,
- НУР — гіпноное,
- А - апное.

На підставі кількості виявлених апное-епізодів за годину сну розраховується індекс АНІ (апное-гіпноное індекс), який класифікує тяжкість синдрому: легка (5–15), середня (15–30) та важка (>30) подій на годину. Ці результати зберігаються в базі даних і відображаються у вигляді часових рядів та інтегрованих метрик, доступних для лікарів і пацієнтів через інтерфейс аналітики. Результати зберігаються у спеціалізованих базах даних, що забезпечує зручний доступ для лікарів і пацієнтів. Ці дані можуть бути представлені у вигляді структурованих звітів або інтегровані в історичні записи пацієнта, дозволяючи відстежувати зміни у його стані здоров'я протягом часу. Для збереження прогнозів використовуються Time-Series бази даних, які оптимізовані для роботи з даними, що мають часову прив'язку, або реляційні бази даних, якщо результати вимагають складнішої структури. Вибір типу бази даних залежить від специфіки даних і цілей аналізу, забезпечуючи ефективне управління результатами та їх доступність для подальшого використання.

3.5.9. Відправка даних на мобільний пристрій

Після формування результатів класифікації, дані відправляються до мобільного додатка через API Gateway. Мобільний додаток отримує ці дані, які

можуть включати інформацію про стан пацієнта або рекомендації щодо подальших дій, і відображає їх користувачеві у зручному форматі. У разі виявлення критичних змін у стані здоров'я система автоматично надсилає сповіщення, забезпечуючи оперативне інформування пацієнта або медичного персоналу. Завдяки використанню подієво-орієнтованої архітектури процес передачі та відображення даних здійснюється автоматично у відповідь на події, такі як завершення аналізу або отримання нових прогнозів. Такий підхід забезпечує швидкість і надійність у реагуванні системи на зміни стану здоров'я пацієнта, що є критично важливим для медичних застосувань.

3.6. Опис архітектури

Архітектура системи (Рис. 3.7.) складається з трьох основних рівнів: периферійного рівня (прикінцеві обчислення), хмарного рівня та рівня мобільного додатка.

На периферійному рівні здійснюється збір даних за допомогою IoT-пристрою ESP32, оснащеного медичними сенсорами. Часткова обробка даних виконується як на самому пристрої, так і на мобільному додатку, що дозволяє зменшити обсяг даних, які передаються далі. Дані передаються на хмарний рівень через API Gateway, який виконує функції маршрутизації, автентифікації та захисту інформації. У хмарній інфраструктурі дані обробляються та аналізуються за допомогою мікросервісів і моделей штучного інтелекту, які забезпечують прогнозування ризиків та формування рекомендацій. Оброблені результати передаються на рівень мобільного додатка, де вони відображаються користувачеві у зручному форматі. Додаток може також надсилати сповіщення у разі критичних змін у стані здоров'я пацієнта, забезпечуючи своєчасне інформування.

- Периферійний рівень (Edge Computing): Дані збираються та частково обробляються на рівні ESP32 і мобільного додатка. Цей рівень забезпечує зниження навантаження на хмарну інфраструктуру та швидшу обробку критично важливих даних.
- Хмарний рівень (Microservices + EDA): Дані передаються через API Gateway до хмарної інфраструктури, де мікросервіси обробляють і зберігають їх у

відповідних базах даних. Тут використовується подієво-орієнтована архітектура для обробки подій у реальному часі та автоматичної активації процесів обробки та аналізу.

- Аналіз і прогнозування (AI): Оброблені дані передаються до сервісу штучного інтелекту для прогнозування стану здоров'я пацієнтів. Це дозволяє системі виявляти потенційні проблеми зі здоров'ям і надавати рекомендації.
- Повернення результатів: Після обробки результати зберігаються та передаються назад на мобільний додаток, де пацієнт або лікар можуть отримати доступ до них у режимі реального часу.

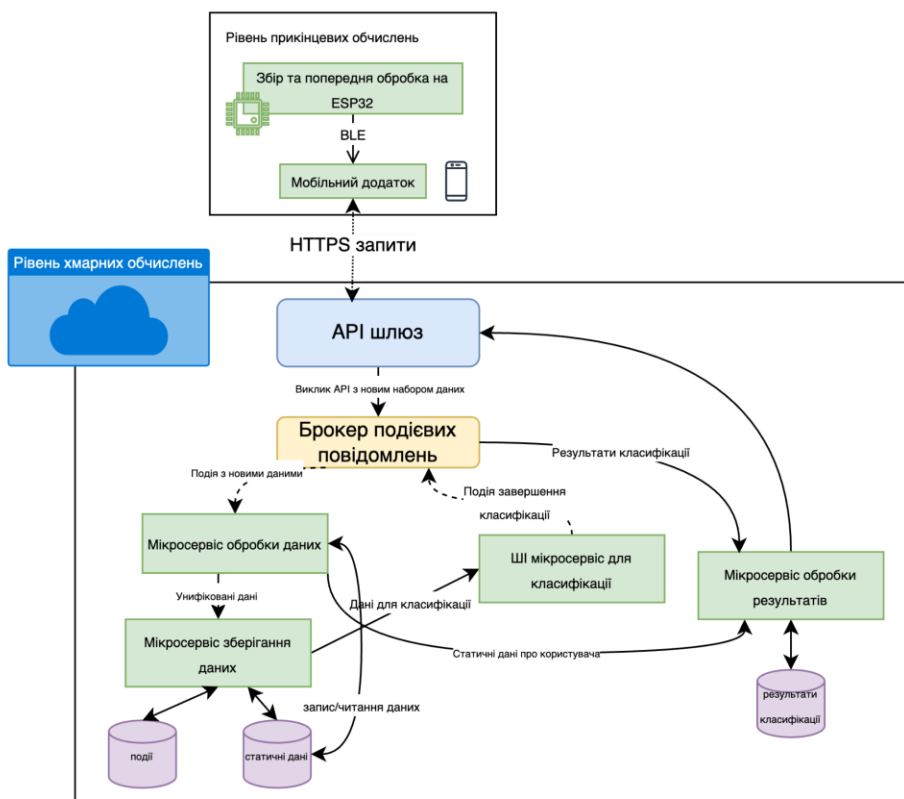


Рис. 3.7. – Архітектура IoT системи для класифікації апное

Розробка системи моніторингу пацієнтів базується на інтеграції технологій edge computing та подієво-орієнтованої архітектури (EDA), що забезпечує високу ефективність і гнучкість у роботі з медичними даними. Завдяки використанню edge computing, дані збираються і проходять попередню обробку безпосередньо на пристрої ESP32, що дозволяє зменшити затримки у передачі та забезпечити швидкий доступ до медичних показників через Bluetooth Low Energy (BLE). Після збору дані передаються на мобільний додаток, який виступає у ролі проміжного елемента між пацієнтом і системою. Це створює безпосередній зв'язок, дозволяючи пацієнтам отримувати інформацію про свій стан у реальному часі та своєчасно реагувати на потенційні загрози здоров'ю. Така архітектура поєднує локальну обробку з масштабною аналітикою у хмарі, забезпечуючи баланс між продуктивністю та зручністю.

Після передачі даних на хмарну платформу, подієво-орієнтована архітектура (EDA) забезпечує їх асинхронну обробку за допомогою різноманітних мікросервісів. Кожен мікросервіс виконує чітко визначені завдання: обробку, зберігання та аналіз отриманої інформації, а також прогнозування можливих захворювань за допомогою моделей штучного інтелекту. EDA дозволяє системі реагувати на кожен новий пакет даних у реальному часі, зменшуючи затримки у процесі обробки та забезпечуючи високу чутливість до змін. Такий підхід робить систему надзвичайно гнучкою та масштабованою, дозволяючи адаптувати її до зростаючих потреб користувачів або збільшення кількості підключених пристроїв. Це особливо важливо для медичних IoT-систем, де точність і швидкість обробки даних відіграють ключову роль у забезпеченні своєчасної допомоги пацієнтам.

Попри значні переваги, реалізація подієво-орієнтованої архітектури в поєднанні з edge computing має свої недоліки, які необхідно враховувати. Складність впровадження такої архітектури може вимагати значних ресурсів, зокрема часу, фінансів та висококваліфікованих спеціалістів. Управління даними на edge-пристроях потребує особливої уваги, щоб уникнути їх втрати або пошкодження, що може суттєво вплинути на точність системи. Хоча edge computing зменшує ризики, пов'язані з передачею даних, він також відкриває нові

вектори потенційних атак, що підвищує вимоги до безпеки. Залежність від стабільного з'єднання між edge-пристроями та хмарними сервісами створює ризик затримок або втрати даних у разі перебоїв у мережі. Збільшення кількості edge-пристроїв може ускладнити управління системою, особливо якщо немає уніфікованих стандартів, які полегшують інтеграцію різних компонентів. Це може вплинути на ефективність системи, підвищуючи складність масштабування та управління. Тому для успішного впровадження архітектури необхідно ретельно планувати заходи з безпеки, стандартизації та управління даними.

Інтеграція edge computing та подієво-орієнтованої архітектури в архітектуру системи не лише оптимізує обробку даних, але й забезпечує високу гнучкість у реагуванні на зміни. Такий підхід робить систему ефективним інструментом для моніторингу пацієнтів у реальному часі, дозволяючи зменшити затримки, підвищити швидкість обробки інформації та забезпечити її своєчасну передачу. Однак, для забезпечення надійності та безпеки роботи системи необхідно враховувати низку викликів, пов'язаних із впровадженням цих технологій. Це включає управління складністю архітектури, мінімізацію ризиків втрати даних та реалізацію надійного захисту від кібератак. Лише ретельне планування та дотримання стандартів можуть гарантувати успішну реалізацію та ефективність системи в умовах зростаючих вимог до обробки медичних даних.

РОЗДІЛ 4. ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ НАВЧАННЯ МОДЕЛЕЙ ДЛЯ КЛАСИФІКАЦІЇ АПНОЕ СНУ

Швидкий розвиток технологій штучного інтелекту суттєво впливає на сучасне суспільство, відкриваючи широкі можливості для впровадження ШІ у різноманітних сферах. Зокрема, машинне навчання, як ключовий компонент ШІ, демонструє виняткову здатність виявляти та класифікувати складні закономірності у великих масивах даних. Ця особливість робить його незамінним інструментом для роботи з великими й складними наборами даних, зокрема у сфері охорони здоров'я, що підтверджують численні попередні дослідження. Застосування методів машинного навчання у медичних та санітарних галузях стрімко поширюється завдяки можливостям аналізувати багатофакторні дані та будувати прогностичні моделі. Особливо актуальним це стає для дослідження складних станів, таких як обструктивне апное сну (ОАС), розвиток якого залежить від багатьох взаємозалежних факторів. У цьому контексті машинне навчання виявляється надзвичайно ефективним інструментом для розробки моделей прогнозування, які можуть забезпечити більш точну діагностику та моніторинг стану пацієнтів [54, 55].

Обструктивне апное сну (ОАС), яке відзначається складною багатофакторністю та різноманітними причинами, стало об'єктом численних досліджень, спрямованих на його фенотипування. У цьому контексті кластеризація, що належить до методів неконтрольованого навчання у машинному навчанні, набула широкого застосування. Її ефективність обумовлена здатністю аналізувати багатовимірні дані без попередніх міток, що дозволяє виявляти приховані закономірності та групи серед пацієнтів [56, 57, 58]. На основі цього підходу, дане дослідження ставить за мету підвищення точності класифікації шляхом інтеграції кластеризації як попереднього етапу перед застосуванням методів контрольованого навчання. Такий підхід дозволяє структурувати дані,

зменшити їхню варіативність і забезпечити більш точні результати класифікації. Це має важливе значення для детального аналізу та розробки ефективних моделей діагностики ОСАС, що орієнтовані на багатофакторний підхід.

4.1 Вибір та реалізація алгоритмів машинного навчання для прогнозування захворювань дихання та апное

Вибір алгоритмів машинного навчання для прогнозування захворювань дихальної системи, таких як обструктивне апное сну, є складним та багатогранним завданням. Цей процес вимагає врахування характеристик даних, таких як їхня структура, обсяг і наявність пропусків, а також вимог до продуктивності моделі, що включають точність, швидкість роботи й здатність працювати в режимі реального часу. Крім того, важливим фактором є інтерпретованість моделі, особливо у сфері охорони здоров'я, де лікарі повинні мати змогу зрозуміти логіку прогнозів для ухвалення клінічних рішень. Також значення мають обчислювальні ресурси, доступні для навчання та впровадження моделі, оскільки складні алгоритми можуть бути неефективними для використання на пристроях з обмеженою потужністю, таких як IoT-пристрої. Для успішного прогнозування ОСАС доцільно розглянути кілька класів алгоритмів, кожен з яких має свої переваги та обмеження. Наприклад, рекурентні нейронні мережі (RNN) є ефективними для аналізу часових рядів, однак можуть бути складними в навчанні. Таким чином, вибір алгоритму повинен базуватися на збалансованому врахуванні специфіки задачі та доступних ресурсів, що дозволить забезпечити оптимальні результати прогнозування.

Рекурентні нейронні мережі (RNN) та їхні модифікації, зокрема Long Short-Term Memory, вважаються одними з найперспективніших підходів для роботи із часовими рядами. Завдяки здатності враховувати довготривалі залежності у даних, ці моделі ідеально підходять для аналізу динамічних сигналів, таких як фотоплетизмографічні сигнали та частота серцевих скорочень. Використання LSTM дозволяє здійснювати точне прогнозування, навіть якщо дані мають складну часову структуру. Коли ж мова йде про обробку сигналів зі складними локальними залежностями, наприклад ФПГ, де ключову інформацію несуть різкий систолічний

фронт, дикротична вирізка, співвідношення ширини та крутизни хвилі й дихальна модуляція амплітуди, ефективним підходом є поєднання згорткових мереж із LSTM. Така комбінація дозволяє CNN виявляти локальні просторові особливості сигналу, тоді як LSTM моделює довготривалі залежності, забезпечуючи високу точність аналізу та прогнозування. Для моделювання залежностей між змінними, такими як ЧСС, SpO_2 та класифікація апное, ефективними є байєсівські мережі. Цей підхід дозволяє будувати ймовірнісні моделі, які враховують невизначеність та взаємозв'язки між змінними, що є важливим для задач медичної діагностики, де дані можуть бути неповними або мати високу варіативність. Таким чином, поєднання LSTM, CNN та байєсівських мереж відкриває нові можливості для створення комплексних моделей, здатних точно прогнозувати ризики захворювань на основі біометричних даних, враховуючи як локальні, так і глобальні залежності у часі.

Для обґрунтованого вибору моделі необхідно порівняти алгоритми за кількома критеріями, такими як точність прогнозу, швидкість навчання та виконання, здатність до узагальнення, інтерпретованість і стійкість до шумів. Точність моделей оцінюється за допомогою метрик, зокрема ROC-AUC, F1-міри, точності (accuracy) та чутливості (recall). Швидкість роботи алгоритму набуває особливої важливості при роботі з великими наборами даних, тому слід враховувати як час навчання, так і ефективність прогнозування. Наприклад, нейронні мережі можуть бути обчислювально затратними, але забезпечують високу точність прогнозування. З іншого боку, методи опорних векторів зазвичай працюють швидше, але можуть бути менш масштабованими і не завжди гарантують таку ж точність на великих обсягах даних. Таким чином, вибір оптимальної моделі залежить від балансу між необхідною точністю, швидкістю та доступними обчислювальними ресурсами для навчання і впровадження моделі.

Стійкість до шуму є критичним аспектом для медичних даних, оскільки вони часто містять артефакти або похибки вимірювань. Наприклад, дерева рішень схильні до перенавчання на зашумлених даних, тоді як ансамблеві методи, такі як Random Forest, краще справляються з цією проблемою завдяки об'єднанню рішень

кількох дерев. У контексті інтерпретованості, методи на основі дерев рішень та SVM із лінійним ядром мають перевагу над нейронними мережами, оскільки останні є менш прозорими для користувачів.

Навчання та тестування обраних моделей виконується з використанням навчального та тестового наборів даних. Попередньо дані поділяються на навчальну (70-80%) і тестову (20-30%) множини. Для кожного суб'єкта були вирізані 30-секундні вікна сигналів HR і SpO₂; до кожного вікна додано вік, стать, зріст, вагу та індекс маси тіла, після чого визначено бінарну мітку «апноє / норма». Щоб виключити перетин даних одного й того самого пацієнта поділ виконано на 80% учасників потрапили до навчальної множини, 20% — до тестової; функція випадкового поділу вхідних даних забезпечила збереження балансу класів. На навчальній частині застосовано стратифіковану п'ятиразову перехресну валідацію; у кожній ітерації чотири підмножини даних використовувалися для навчання двохшарової LSTM-мережі з 64 елементами пам'яті та коефіцієнтом dropout 0,2, а п'ята підмножина — для валідації.

Функція втрат є ключовим компонентом при навчанні моделей машинного навчання і визначається залежно від типу задачі. У задачах класифікації зазвичай використовуються функції, такі як крос-ентропія або бінарна логістична втрата, які ефективно мінімізують помилки класифікації. Для задач регресії основними метриками оцінювання помилки є середньоквадратична помилка (MSE) або середньоабсолютна помилка (MAE), що дозволяють оцінити різницю між передбаченими та реальними значеннями. Для задач, пов'язаних із часовими рядами, наприклад, при аналізі біометричних сигналів, можуть застосовуватися спеціалізовані метрики. Однією з таких метрик є динамічне викривлення часу (англ. Dynamic Time Warping, DTW), яка враховує зміщення у часі між передбаченим і реальним сигналами. Це дозволяє більш точно оцінити відповідність моделей часових даних, що є критично важливим у медичних додатках, де навіть незначні зміни у сигналі можуть мати суттєве значення для клінічних висновків.

Перед побудовою матриці похибок дослідник спершу очистив і сегментував сигнали, обчислив часові, частотні та морфологічні ознаки варіабельності серцевого ритму (BCR), ФПГ і сатурація разом із демографічними даними, а потім за допомогою статистичних тестів і перестановок важливості ознак відібрав найінформативніші показники. Це дозволяє ідентифікувати фактори, які найбільше впливають на класифікацію апное. Такий підхід сприяє не лише підвищенню точності моделі, але й забезпечує її інтерпретованість, що має вирішальне значення для клінічного застосування. Результати цього аналізу допомагають медикам краще зрозуміти ключові показники, які впливають на стан пацієнтів, і адаптувати відповідні стратегії лікування.

Модель LSTM є одним із найперспективніших інструментів для аналізу часових рядів завдяки своїй здатності враховувати довготривалі залежності у даних. У дослідженні кожна послідовність сигналів, таких як ФПГ, ЧСС і сатурація, представляється у вигляді векторів із часовою розміткою. Вікно часових рядів, яке використовується як вхідний сигнал для моделі, було підібрано емпірично і зазвичай становить 30 секунд. Для забезпечення коректного навчання всі значення сигналів перед подачею на модель нормалізувалися до діапазону $[0, 1]$. Пропущені дані заповнювалися методами інтерполяції або за допомогою середнього значення, щоб зберегти цілісність часових рядів. Навчання моделі проводилося послідовно, з використанням рекурентних блоків LSTM, які забезпечують збереження та аналіз тимчасових залежностей у даних. Функція втрат була вибрана залежно від задачі: для задач класифікації використовувалася крос-ентропія, тоді як для регресійних задач застосовувалася середньоквадратична помилка (MSE). Як оптимізатор було обрано метод Adam (англ. Adaptive Moment Estimation) об'єднує переваги стохастичного градієнтного спуску з моментом та адаптивного регулювання швидкості навчання, підтримуючи окремі коефіцієнти навчання для кожного параметра мережі. На кожному кроці Adam оцінює перший момент (середнє значення градієнта) і другий момент (норму градієнта) із застосуванням експоненційного згладжування, що дозволяє йому швидко досягати збіжності навіть у задачах із високою розмірністю та неконстантною масштабованістю ознак

. Adam також автоматично коригує швидкість навчання, що робить його стійким до вибору початкових гіперпараметрів (швидкість навчання 0,001; коефіцієнт згладжування першого моменту 0,9; коефіцієнт згладжування другого моменту 0,999; $\epsilon = 1 \times 10^{-8}$) і дозволяє досягати високої точності у глибоких моделях.

Архітектура нейронної мережі Transformer, яка ґрунтується на механізмі уваги, представляють перспективний підхід до аналізу часових рядів, що відрізняється своєю здатністю враховувати залежності між усіма елементами послідовності одночасно. На відміну від моделей LSTM, трансформери не мають обмежень на довжину цих залежностей, що робить їх ефективними для роботи з довготривалими часовими рядами. У процесі підготовки до навчання послідовності ФПГ, ЧСС та SpO₂ розбиваються на 10-секундні сегменти (1000 відліків при 100 Гц) із кроком 2 с (рис. 4.1.), після чого для кожного сегмента миттєва ЧСС обчислюється «від удару до удару» у 2-секундному вікні (3–5 піків при ЧСС 60–90 уд/хв), а SpO₂ визначається у ковзному 5-секундному вікні з кроком 1 с (5 оцінок за сегмент). До кожного вхідного вектору додаються синусоїдальні позиційні кодування для збереження порядку в послідовності. Ці кодування дозволяють зберігати інформацію про порядок елементів у послідовності, що є критично важливим для аналізу часових даних. Вхідні дані також проходять етапи нормалізації та дискретизації, щоб забезпечити узгодженість та зменшити шум. Навчання трансформерів здійснюється з використанням оптимізатора AdamW, який поєднує адаптивне регулювання швидкості навчання з відокремленим розрахунком L2-регуляризації, що в свою чергу поєднує адаптивне регулювання швидкості навчання з відокремленим розрахунком згасання ваг. Завдяки цьому він стабільніше контролює величину ваг, знижує ризик перенавчання на великих моделях і прискорює збіжність. Функція втрат обирається залежно від задачі: для класифікації застосовується крос-ентропія, тоді як для регресії — середньоквадратична помилка (MSE). Завдяки своїй здатності паралельно обробляти залежності в усій послідовності, трансформери відкривають нові можливості для точного та швидкого аналізу медичних часових рядів.



Рисунок 4.1. Сегментація ФПГ-сигналу для розрахунку ЧСС Та SpO₂

Модель 1D CNN пропонує ефективний підхід до аналізу необроблених сигналів, таких як PPG, шляхом автоматичного виділення релевантних особливостей. Вхідні дані подаються у вигляді багатоканальних послідовностей, де кожен канал відповідає конкретному параметру, наприклад, частоті серцевих скорочень (HR) або сатурації киснем (SO₂). Основною перевагою 1D CNN є здатність автоматично виявляти локальні залежності у сигналах завдяки застосуванню згорткових шарів, які аналізують невеликі фрагменти даних. Цей підхід дозволяє виявляти важливі патерни, характерні для фізіологічних сигналів. Пулінг-операції (наприклад, max-pooling або average-pooling) додатково зменшують розмірність проміжних представлень, усуваючи надлишкову інформацію та підвищуючи обчислювальну ефективність моделі. Для навчання моделі використовується функція втрат, що відповідає задачі: крос-ентропія застосовується у випадках класифікації, наприклад, для класифікації апное, тоді як для регресії, яка вимагає точного прогнозування фізіологічних параметрів, використовується середньоквадратична помилка. Модель 1D CNN підходить для аналізу біометричних даних завдяки здатності обробляти сигнали без необхідності попередньої ручної інженерії ознак, що особливо важливо для складних та багатовимірних медичних даних.

Для забезпечення коректного порівняння моделей машинного навчання їх продуктивність оцінюється за однакових умов із використанням спільного підходу до підготовки даних. Усі моделі отримують однаковий набір ознак, який створюється на основі ретельно очищених і нормалізованих даних, щоб уникнути

упередженості через різницю у вхідних даних. З метою забезпечення надійності результатів використовується крос-валідація, наприклад, із п'ятиразовим розбиттям даних. Цей підхід допомагає оцінити стабільність моделі та її здатність до узагальнення на різних підмножинах даних. Додатково аналізуються обчислювальні ресурси, включаючи час навчання, використання оперативної пам'яті та енергоспоживання для кожної моделі. Це особливо важливо для складних архітектур, таких як трансформери та LSTM, які можуть вимагати значних ресурсів для обробки великих обсягів даних.

4.1.1. Опис репозиторію цифрових сигналів фотоплетизмограмм для розробки методів діагностики апное сну

Досліджуваний масив даних сформовано з кількох відкритих клінічних баз PhysioNet. До нього увійшли:

- BIDMC PPG and Respiration Dataset – 53 восьмихвилинні записи PPG (125 Гц) критично хворих пацієнтів від Beth Israel Deaconess Medical Center з ручними анотаціями дихальних актів [38];
- St Vincent's UCD Sleep Apnea Database (UCDDB) – нічні полісомнографічні дослідження з багатоканальними сигналами дихання та SpO₂ [67];
- Apnea-ECG Database – 70 записів ЕКГ тривалістю 7–10 год (100 Гц) з хвилинними мітками наявності/відсутності апное [68];
- MIT-BIH Polysomnographic Database (SLPDB) – повноцінні нічні PSG-записи пацієнтів із обструктивним апное (ЕЕГ, ЕКГ, дихальні канали, SpO₂);

Комбінування цих джерел забезпечує репрезентативний набір AGU-, респіраторних та насичених анотаціями даних для тренування й перевірки моделі класифікації апное. Цей набір даних є унікальним ресурсом для аналізу дихальних процесів, оскільки включає фізіологічні сигнали, такі як сигнали ФПГ та ЕКГ, а також численні фізіологічні параметри й демографічні характеристики пацієнтів. Дані використовувалися для розробки, навчання та оцінки алгоритмів машинного

навчання, спрямованих насамперед на виявлення та класифікацію епізодів порушень дихання (гіпопное, обструктивне, центральне та змішане апное), а також на розрахунок таких ключових параметрів, як частота дихання (RR).

У датасеті BIDMC PPG & Respiration Dataset (53 записи \times 8 хв) кожен файл містить фотоплетизмограму, імпедансний респіраторний канал і п'ятиканальний ЕКГ, усі оцифровані з частотою 125 Гц. Записи отримано від критично хворих пацієнтів Beth Israel Deaconess Medical Center під час госпітального спостереження. Кожен файл містить ФПГ, імпедансний респіраторний сигнал і п'ятиканальний ЕКГ, оцифровані з частотою 125 Гц, а також щосекундні агреговані показники ЧСС, частота дихання та сатурації. Два експерти вручну позначили кожен окремий вдих і видих на основі імпедансного сигналу; епізоди апное чи гіпопное в цьому наборі не анотовано. Датасет застосовується для тестування алгоритмів оцінки частоти дихання за ФПГ і формує «нормальний» фон без подій апное.

Натомість у наборі даних “St Vincent’s / UCD Sleep Apnea Database” (25 нічних записів полісомнографії) містяться повнонічні записи ($\approx 6\text{--}8$ год на запис) охоплюють оро-назальний потік, грудне й абдомінальне зусилля, сатурацію (1 Гц) та триканальний ЕКГ (128 Гц). Анотаціями відмічено ≈ 6280 гіпопное, ≈ 4050 обструктивних, ≈ 690 центральних і ≈ 480 змішаних апное; середня тривалість події ≈ 12 с. Ці дані відображають реальні нічні умови й забезпечують генералізацію моделі.

У наборі даних “Apnea-ECG Database” (70 записів по 7–10 год) міститься сигнал одноканального ЕКГ із частотою 100 Гц. Для 35 навчальних-файлів (a01–a20, b01–b05, c01–c10) наявні хвилинні анотації «апное/норма» без поділу за типами; в свою чергу 35 файлів з тестовими даними (x01–x35) анотацій не мають. Сумарно в навчальній частині позначено ≈ 16000 хвилин апное ($\approx 3,5$ год на запис у середньому), що використовують як еталон для валідації детекції епізодів порушень дихання й розрахунку АНІ. Вісім записів (a01–a04, b01, c01–c03) додатково містять чотири респіраторні сигнали (грудне та абдомінальне зусилля, оро-назальний потік) і сатурацію, але анотації залишаються лише хвилинними. Таким чином, база служить незалежним бінарним референтом «апное/норма» для

перевірки узгодженості виходу LSTM-моделі з офіційним АНІ; диференціювати типи апное за цим набором неможливо.

Також набір даних “MIT-BIH Polysomnographic Database” (SLPDB) містить 18 нічних полісомнографічних записів тривалістю від 1 год 17 хв до 6 год 30 хв, отриманих від 16 чоловіків 32–56 р. (середньої маси тіла 119 кг). Кожен файл містить високочастотні канали (до 250 Гц): електрокардіограму, інвазивний артеріальний тиск, два відведення електроенцефалографії, сигнал назального термістора (дихальний потік), а також у більшості випадків грудну або абдомінальну плетизмографію, оксиметрію, електроокулографію та підборідну електроміографію. Анотації у файлах формату .st задають 30-секундні інтервали з позначками стадій сну та типів подій дихання: OA/X — обструктивне апное, CA — центральне, HA/H — гіпопное, а також змішані епізоди й арозали. Завдяки широкому діапазону АНІ (від <1 до >100 подій/год) SLPDB забезпечує повний спектр тяжкості порушень дихання й слугує основою для навчання й перевірки LSTM-класифікатора класів Normal, HYP, O, C, M та розрахунку індексу АНІ.

Записи з баз даних (табл 4.1.) сегментуються 30-секундними вікнами з перекриттям 50 %. Для кожного сегмента присвоюється мітка події за більшістю анотацій у межах вікна та обчислюється локальний АНІ. У підсумку отримано ≈ 110 тис. сегментів (≈ 45 % Normal; інші розподілені O:HYP:C:M ≈ 38 %:45 %:10 %:7 %). Такий балансований корпус забезпечує повноцінне навчання та валідацію LSTM-моделі.

Таблиця 4.1. Приклади сигналів, склад та розподіл епізодів апное у використовуваних базах

| Набір даних | Канали (частота) | К-ть записів із апное | Підраховані епізоди* | Розподіл типом** | Середній/діапазон АНІ |
|-------------|---|-----------------------|----------------------|--------------------------------|-----------------------|
| UCDDB | оро-назальний потік (100 Гц), грудне / черевне зусилля (10 Гц), SpO ₂ (1 Гц), 3-канальний ЕКГ (128 Гц) | 25 / 25 | ≈ 11500 | HYP 6280; O 4050; C 690; M 480 | 5 – 82 подій/год |

| | | | | | |
|-------------------------|---|--|---|-----------------------------------|---|
| SLPDB | ЕКГ, артер. тиск, 2 ЕЕГ, термісторний потік, індукц. плетизмографія, SpO ₂ (≤ 250 Гц) | 17 / 18*** | 8151 | НУР 4240; О 3132; С 512; М 267 | 0.7 – 100.8; 1 норм., 1 легке, 5 помірних, 11 тяжких |
| Apnea-ECG (learning) | одноканал. ЕКГ (100 Гц) \pm Resp/SpO ₂ у 8 файлах | 35 / 35 | ≈ 16000 хв «апное» (без поділу) | — | 14 – 82 подій/год (за файлами a01– c10) |
| BIDMC | PPG (Red/IR), імпедансний подих, 5-кан. ЕКГ (125 Гц) | 0 / 53 (лише нормальні дихальні цикли) | — | — | — |

* Кількість подій підсумовано за еталонними *.st / *.apn файлами; для Apnea-ECG наводиться сумарна кількість хвилин, позначених «апное».

** О — обструктивне, С — центральне, М — змішане апное, НУР — гіпопное.

*** Файл slp67x має АНІ 0.7 → використовується як еталон «норма».

Загальна база даних охоплює 78 нічних записів, із яких 77 містять принаймні один епізод апное чи гіпопное, тоді як BIDMC слугує «чистим» фоном без порушень дихання. Сукупно анотовано понад 19600 подій: близько 10500 гіпопное, 7200 обструктивних і 1200 центральних або змішаних епізодів, що формує збалансовану вибірку для навчання LSTM-класифікатора п'яти класів (Normal, НУР, О, С, М). Діапазон індексу АНІ перекриває всі ступені тяжкості — від <5 до >100 подій на годину — тож модель не лише виявляє апное, а й оцінює його стадію; багатоканальні PSG-записи з точними еталонними анотаціями забезпечують повноцінне тренування, а BIDMC дає змогу перевірити рівень хибнопозитивних спрацьовувань на нормальному сигналі.

4.1.2. Підготовка даних для застосування методів глибокого навчання до задачі класифікації сигналів фотоплетизмограмм для виявлення апное сну

Процес підготовки датасету для навчання моделей штучного інтелекту є критично важливим етапом, який безпосередньо впливає на точність і ефективність прогнозування. У цьому дослідженні об'єднання даних із множини файлів дозволило уніфікувати структуру набору, створивши єдиний інтегрований датасет,

придатний для подальшого аналізу. Це уніфікування забезпечує послідовність у форматах даних, спрощуючи їх обробку та інтеграцію. Реалізація автоматизованого процесу об'єднання 53 індивідуальних записів із фізіологічними даними через програмний код забезпечує систематичний і відтворювальний підхід. Такий підхід значно зменшує ймовірність людських помилок і дозволяє працювати з даними ефективно навіть при збільшенні обсягу записів. Уніфікований набір даних стає основою для коректного тренування моделей, оскільки забезпечує стандартизовану форму вхідної інформації та підтримує можливість масштабування для більш складних і великих аналізів.

Початкові дані представлені у вигляді окремих файлів, що містять різні типи інформації: фізіологічні параметри, такі як ЧСС та сатурація, а також сигнали та мета інформацію про пацієнтів, наприклад, вік і стать. Файли з числовими даними зчитуються з відповідних CSV-файлів, тоді як текстові файли забезпечують доступ до демографічної інформації. Під час обробки кожного запису реалізовано інтеграцію цих джерел інформації, що дозволяє забезпечити узгодженість між різними рівнями даних. Так використання 0 або середнього значення замість пропущених даних можуть спотворювати розподіл: нуль зміщує дані, а середнє чутливе до поодиноких екстремальних значень. Тому пропущені відліки потрібно замінити медіаною відповідного параметра всередині того самого 10-секундного (або нічного) вікна, оскільки медіана стійка до викидів і краще відображає типові значення. Якщо провал у сигналі триває менше ніж одну секунду, перед заповненням медіаною застосовуємо лінійну інтерполяцію між сусідніми доступними точками, а після заміщення виконуються стандартні кроки масштабування. Цей крок має критичне значення, оскільки гарантує цілісність даних і запобігає помилкам під час навчання моделей машинного навчання. Такий підхід забезпечує стандартизований та уніфікований набір даних, що є ключовим для коректного аналізу. Це дозволяє створити надійну базу для подальшого тренування моделей, підвищуючи їх точність і стійкість до впливу шуму або прогалин у даних.

Крім обробки пропущених даних, інтеграція фізіологічних сигналів, таких як ФПГ, із числовими параметрами, такими як ЧСС і SpO_2 , формує багатовимірний набір ознак. Цей підхід забезпечує гнучкість у використанні даних, дозволяючи працювати як із часовими рядами, так і з агрегованими характеристиками, що є особливо корисним для побудови LSTM моделей машинного навчання. Ключовим аспектом цього процесу є узгодження розмірності між сигналами та числовими даними, оскільки різна кількість записів у цих наборах може вплинути на коректність аналізу. Для досягнення цієї узгодженості використовуються техніки, такі як обрізання надлишкових записів у сигналах до довжини, відповідної числовим параметрам. Такий метод підготовки створює структурований і уніфікований вхідний формат, який сприяє підвищенню продуктивності моделей за рахунок одночасного використання інформації з різних джерел, зберігаючи часові залежності та агреговані ознаки.

Для інтеграції метаінформації, такої як вік і стать, дані з текстових файлів проходять етап нормалізації. Вік трансформується у числові значення, де категорії, такі як "90+" або "NaN", перетворюються у медіанне значення. Це дозволяє враховувати старших пацієнтів і обробляти пропущені значення без втрати інформації. Стать пацієнтів також перетворюється у числові значення: "М" кодується як 1, а "Ж" як 2. Такий підхід дозволяє легко інтегрувати ці ознаки в алгоритми машинного навчання, які працюють із числовими даними. На етапі об'єднання всі джерела даних інтегруються в єдиний DataFrame, який поєднує метаінформацію, числові параметри та часові ряди. Цей процес побудований так, щоб уникнути втрати інформації, забезпечити узгодженість між записами та підтримувати однаковий формат даних для кожного пацієнта. Завдяки цьому повний набір даних має структуру, придатну для різних алгоритмів машинного навчання. Заключним етапом є збереження уніфікованого набору даних у форматі CSV, що слугує основою для подальшого аналізу, тренування моделей та валідації. Такий підхід забезпечує легкість доступу, повторюваність експериментів та узгодженість під час роботи з даними.

Підготовлений набір даних оптимізовано для тренування моделей штучного інтелекту, оскільки він поєднує інформацію з різнорідних джерел, ефективно обробляє пропущені значення та забезпечує узгодженість між усіма записами. У результаті створюється структурований і збалансований набір даних: після розбиття всіх записів на 30-секундні вікна отримано близько 110 000 сегментів із таким початковим розподілом: Нормальний ≈ 50000 , НУР ≈ 25000 , О ≈ 20000 , С ≈ 10000 і М ≈ 5000 . Щоб запобігти перевазі найбільшого класу, з нього випадковим чином видалили 25000 сегментів. Для рідкісних класів С і М штучно збільшили вибірку: до існуючих додано нові сегменти, отримані віддзеркаленням сигналу та випадковим часовим зсувом, поки кожен із п'яти класів не містив по 25000 прикладів. У підсумку сформовано рівноважний набір із 125000 сегментів, де всі класи представлені однаковою кількістю спостережень, який включає числові, категоріальні та часові ознаки. Така універсальна структура відкриває можливості для використання різних підходів до аналізу, від класичних алгоритмів машинного навчання до глибоких нейронних мереж. Попередня обробка даних (заміна пропусків, нормалізація, вирівнювання класів) створює єдину узгоджену структуру, на якій можна формувати набори ознак, оцінювати їхню інформативність і порівнювати продуктивність різних моделей класифікації дихальних подій.

Додано примітку [1]: а як дані було збалансовано по різних класам? У тесті відсутня інформація, скільки сигналів в базі даних відносилось до здорових людей, скільки мали проблеми саме з диханням, а конкретно з апное. Потім ви мали їх балансувати, щоб в різних класах була приблизно однакова кількість спостережень

4.2. Порівняння результатів машинного навчання для різних архітектур нейронних мереж

Оцінюючи модель, важливо розглядати всі три метрики у сукупності. MSE допомагає визначити моделі, що погано працюють із великими похибками, тоді як MAE забезпечує більш інтуїтивну оцінку. R^2 , у свою чергу, дозволяє оцінити, наскільки добре модель пояснює варіацію у даних. Якщо MSE значно більше за MAE, це може свідчити про те, що модель допускає великі похибки в окремих прогнозах. З іншого боку, якщо R^2 негативне або близьке до нуля, це може свідчити про недостатню продуктивність моделі або наявність надмірного шуму в даних.

Порівнюючи результати моделей, слід зосередитися на балансі між навчальним і тестовим наборами. Якщо модель демонструє значно кращі результати на навчальних даних, ніж на тестових, це вказує на перенавчання. У такому випадку необхідно застосовувати методи регуляризації або зменшувати складність моделі. Якщо результати однаково погані на обох наборах, це може свідчити про недонавчання, і потрібно збільшити складність моделі або вдосконалити набір даних.

Окрім метрик, для вибору моделі важливо використовувати візуалізації, такі як графіки порівняння реальних і передбачених значень, гістограми залишків і scatter plot. Такі інструменти дозволяють краще зрозуміти природу помилок моделі та зробити більш обґрунтований вибір. У підсумку, найкращою є модель, яка має низькі значення MSE і MAE, високе R^2 , стабільні результати на обох наборах даних і демонструє рівномірний розподіл залишків.

Таб. 3.2. Метрики навчального датасету

| | MSE | MAE | R^2 |
|--------------|---------|--------|--------|
| LSTM | 2.3182 | 0.9274 | 0.8294 |
| Transformers | 12.9966 | 2.4641 | 0.0437 |
| 1D-CNN | 5.3258 | 1.5917 | 0.6034 |

Таб. 3.3. Метрики тестового датасету

| | MSE | MAE | R ² |
|--------------|---------|--------|----------------|
| LSTM | 2.4541 | 0.9514 | 0.8213 |
| Transformers | 13.0859 | 2.4399 | 0.0470 |
| 1D-CNN | 5.2760 | 1.5918 | 0.6326 |

Аналіз метрик оцінювання моделей LSTM, Transformers і 1D-CNN демонструє їхню відносну ефективність у задачі прогнозування. Середньоквадратична помилка (MSE), яка відображає середній квадрат різниць між реальними та прогнозованими значеннями, дозволяє оцінити точність моделей. Модель LSTM показує найнижчі значення MSE (2.3182 для навчального набору та 2.4541 для тестового), що свідчить про її здатність мінімізувати похибки та забезпечувати точні прогнози. У порівнянні, модель 1D-CNN демонструє помітно вищі значення MSE (5.3258 для навчального і 5.2760 для тестового наборів), що вказує на її обмежену ефективність у цій задачі. Моделі Transformers, зі значеннями MSE понад 12, демонструють найгіршу продуктивність, що свідчить про суттєві похибки в їхніх прогнозах і вказує на необхідність додаткової оптимізації або застосування іншого підходу для цієї задачі. Таким чином, LSTM виявляється найбільш ефективною моделлю, забезпечуючи найменшу похибку серед усіх розглянутих архітектур.

Середньоабсолютна похибка (Mean Absolute Error, MAE), яка вимірює середню абсолютну різницю між реальними і прогнозованими значеннями, підтверджує перевагу моделі LSTM. Ця архітектура демонструє MAE 0.9274 для навчального набору і 0.9514 для тестового, що свідчить про її стабільність і високу точність у прогнозуванні. Модель 1D-CNN показує середні результати із значеннями MAE 1.5917 для навчального і 1.5918 для тестового наборів, посідаючи проміжну позицію між іншими архітектурами. Натомість моделі Transformers

Додано примітку [2]: для якої цільової змінної? Частота дихання? Але де зв'язок саме з апное?

демонструють найвищі показники похибок: MAE становить 2.4641 для навчального і 2.4399 для тестового наборів. Це вказує на їхню обмежену ефективність у цій задачі без додаткової оптимізації. Отже, MAE підтверджує, що LSTM залишається найбільш точною і стабільною моделлю серед усіх розглянутих архітектур.

Оцінюючи R^2 , який відображає здатність моделі пояснювати варіацію у даних, LSTM знову демонструє найкращі результати, із значеннями 0.8294 на навчальному і 0.8213 на тестовому наборах. Це свідчить про те, що модель успішно захоплює закономірності в даних і забезпечує точні передбачення. 1D-CNN, хоча й показує помірні результати (0.6034 і 0.6326), все ж поступається LSTM. Найгіршими є результати Transformers із майже нульовим R^2 (0.0437 і 0.0470), що вказує на їхню неспроможність пояснювати варіації в цільовій змінній.

Аналіз графіків результатів роботи моделей LSTM, 1D-CNN та Transformers дозволяє оцінити їхню здатність прогнозувати параметр RESP і порівняти якість передбачень.

На графіку реальних і передбачених значень для моделі LSTM (Рис. 4.2.) синя крива, що представляє реальні значення RESP, майже повністю збігається з помаранчевою кривою, яка відображає передбачені значення. Це свідчить про високу точність моделі. Відхилення між реальними та передбаченими значеннями є мінімальними і спостерігаються лише на окремих ділянках, що підкреслює здатність моделі захоплювати основні закономірності в даних. Розподіл залишків для LSTM (Рис. 4.3.) має чітку симетрію навколо нуля і позбавлений значних відхилень у хвостах, що вказує на відсутність систематичних похибок та високу стабільність моделі.

Додано примітку [3]: для якої задачі?

Додано примітку [4]: що саме є цільовою змінною? За якими іншими показниками вона передбачається?

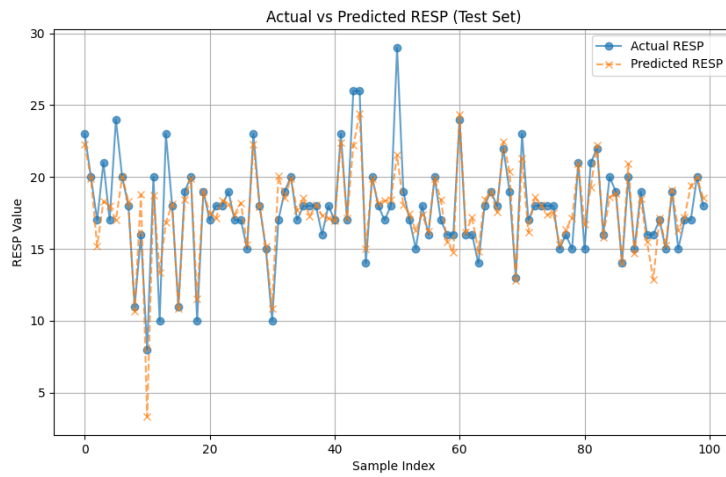


Рис. 4.2. – Графік реальних і передбачених значень для моделі LSTM

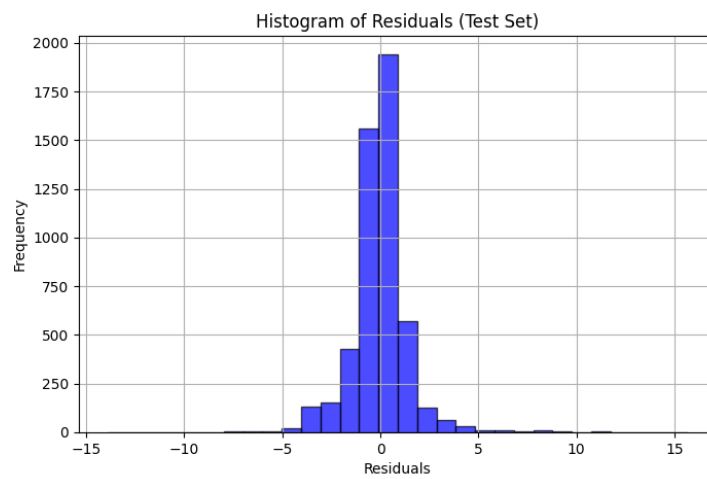


Рис. 4.3. – Розподіл залишків для LSTM

Модель 1D-CNN також демонструє прийнятну продуктивність, але поступається LSTM. На графіку реальних і передбачених значень (Рис. 4.4.) передбачення моделі є досить близькими до реальних значень, хоча на певних ділянках спостерігаються помітні відхилення. Це вказує на те, що модель 1D-CNN не завжди коректно захоплює складні залежності в даних. Розподіл залишків (Рис. 4.5.) для цієї моделі близький до симетричного і зосереджений навколо нуля, що свідчить про стабільність прогнозів. Проте ширина розподілу залишків є більшою, ніж у LSTM, а scatter plot (Рис. 4.6..) виявляє значне розсіювання точок навколо діагоналі, особливо для крайніх значень RESP. Таким чином, 1D-CNN забезпечує задовільну точність, але не досягає рівня LSTM.

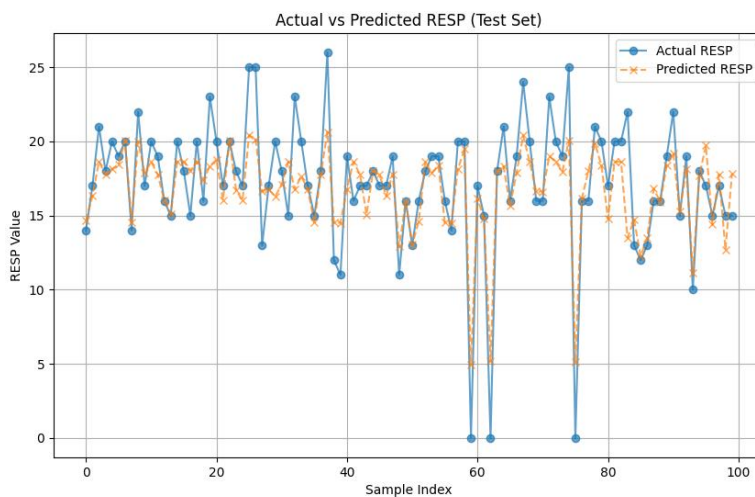


Рис. 4.4. – Графік реальних і передбачених значень 1D-CNN

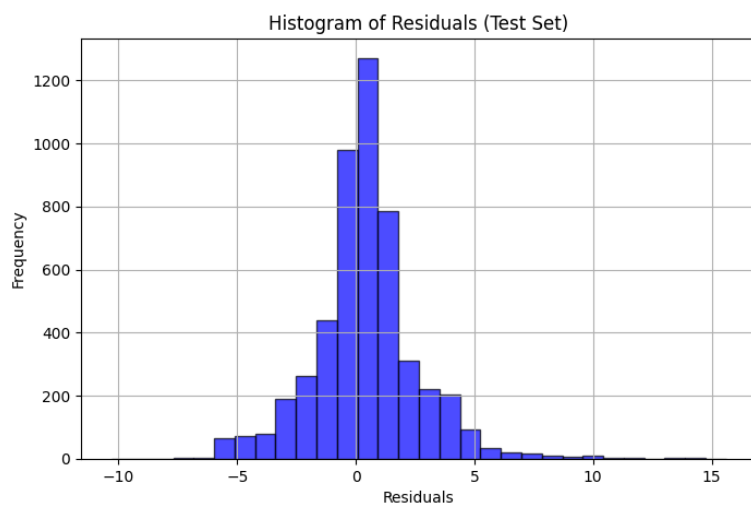


Рис. 4.5. – Розподіл залишків для моделі 1D-CNN

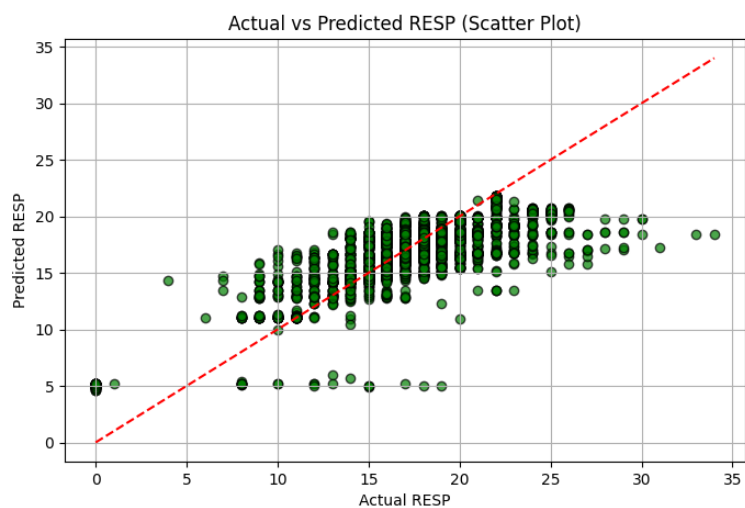


Рис. 4.6. – Ширина розподілу залишків моделі 1D-CNN

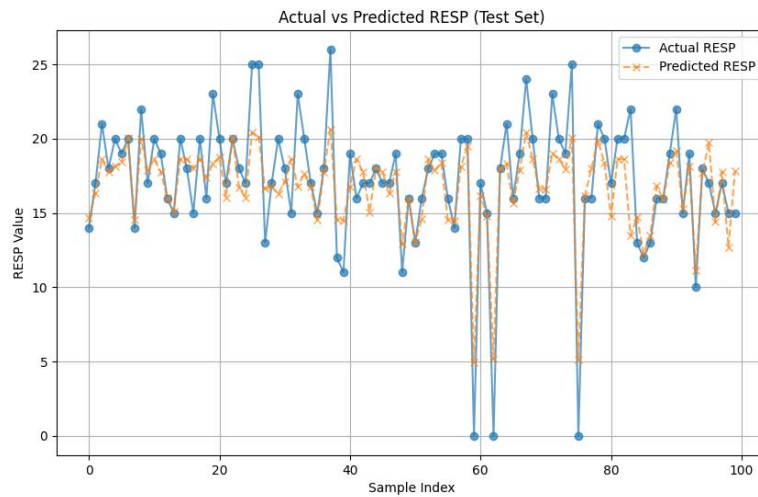


Рис. 4.7. – Графік реальних і передбачених значень 1D-CNN

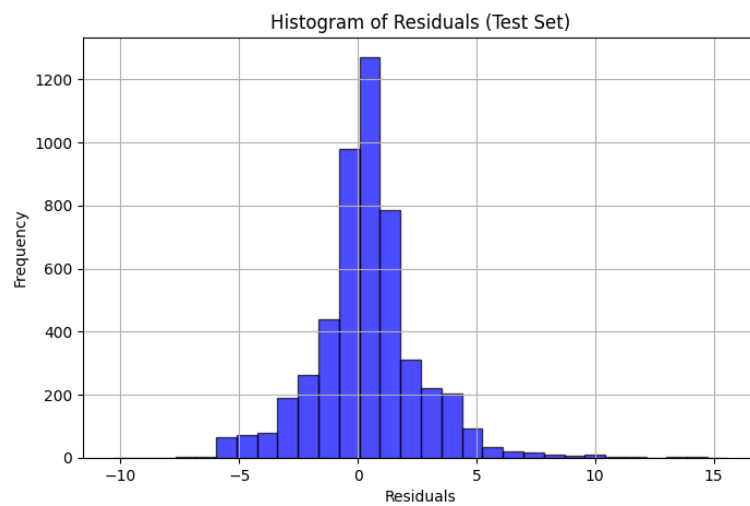


Рис. 4.8. – Розподіл залишків для моделі 1D-CNN

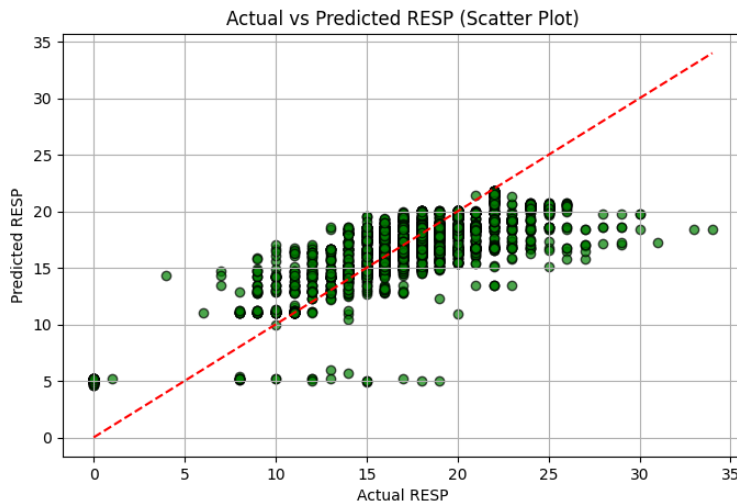


Рис. 4.9 Ширина розподілу залишків моделі 1D-CNN

Натомість модель Transformers демонструє найгірші результати. Графік реальних і передбачених значень (Рис. 4.7.) виявляє значні розходження між синьою та помаранчевою кривими, що свідчить про невідповідність моделі до задачі. Передбачення моделі суттєво відхиляються від реальних значень, що може бути наслідком як архітектурних обмежень трансформерів для цього типу даних, так і недостатнього налаштування моделі. Гістограма залишків (Рис. 4.8.) підтверджує цей висновок: розподіл залишків є асиметричним, широким і зсунутий відносно нуля, що свідчить про наявність систематичних похибок. Діаграма розсіювання (Рис. 4.9.) демонструє, що точки практично не лежать уздовж діагональної лінії, а їх розсіювання є значним. Це вказує на низьку якість прогнозів і недоцільність використання моделі Transformers для цієї задачі.

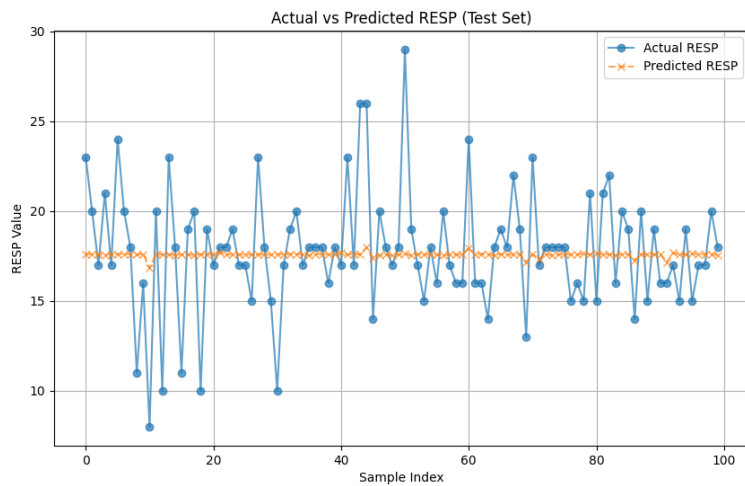


Рис. 4.10. – Графік реальних і передбачених значень моделі Transformers

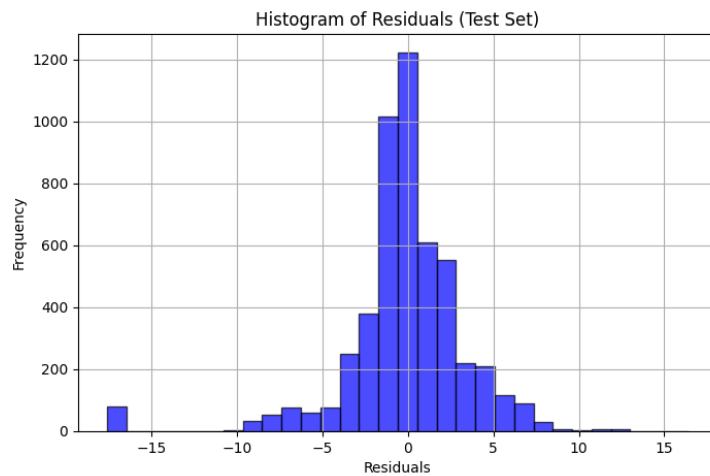


Рис. 4.11. – Гістограма залишків для моделі Transformers

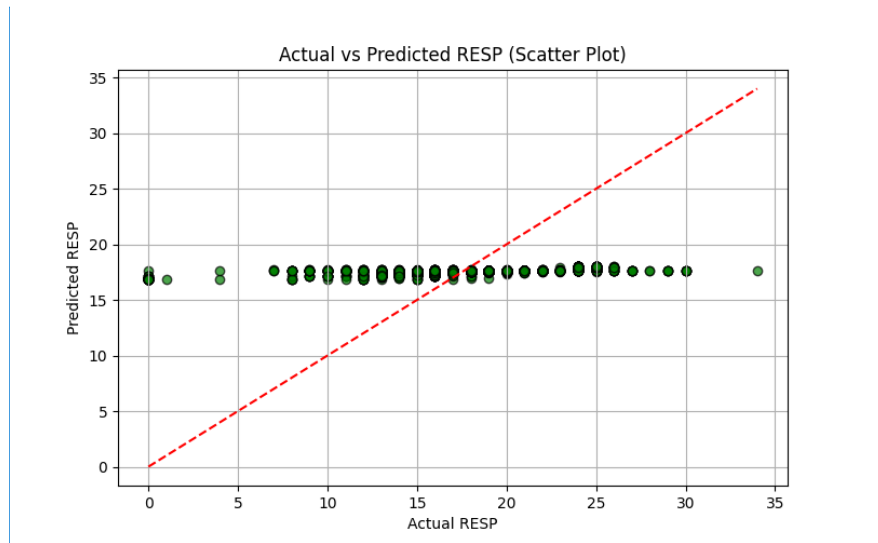


Рис. 4.12. – Ширина розподілу залишків моделі Transformers

Додано примітку [5]: не працює взагалі

Результати графічного аналізу чітко підтверджують перевагу моделі LSTM над іншими алгоритмами в контексті розв'язання задачі. LSTM демонструє найвищий рівень точності, стабільності та здатності до узагальнення даних, що робить її найбільш підходящим вибором для виконання поставленого завдання. Хоча модель 1D-CNN показує прийнятні результати, її продуктивність суттєво поступається LSTM. Водночас модель Transformers виявляється найменш ефективною через значні похибки, що свідчить про її невідповідність до специфіки задачі або недостатню оптимізацію параметрів.

Аналіз метрик MSE, MAE та R^2 , а також їхня стабільність між навчальним і тестовим наборами, підтверджують, що модель LSTM забезпечує мінімізацію похибок, стабільність прогнозів та високу здатність до узагальнення. Модель 1D-CNN може бути використана як резервний варіант, але її продуктивність значно поступається LSTM. Натомість модель Transformers, через низьку ефективність і невідповідність специфіці задачі, є найменш придатною для використання в цьому контексті.

4.3. Навчання моделей на підготовлених даних, оптимізація гіперпараметрів, оцінка якості моделей за допомогою метрик

У задачі прогнозування параметра RESP було використано три підходи: моделі LSTM, 1D-CNN та Transformers. Для забезпечення коректності та узгодженості порівняння, дані були попередньо оброблені відповідно до вимог кожної моделі. Зокрема, кожен набір даних нормалізовано за допомогою StandardScaler, що дозволило звести значення змінних до єдиного масштабу, мінімізуючи вплив різниці в масштабах параметрів на навчання моделей. Такий підхід забезпечив стабільність оптимізації й ефективність роботи алгоритмів. Щоб гарантувати справедливість оцінки моделей, вони були навчені на однаковому розподілі навчального і тестового наборів даних. Для цього було використано стандартне співвідношення 80:20 між навчальними та тестовими даними. Це дозволило оцінити здатність моделей до узагальнення, зберігаючи однакові умови для кожного алгоритму.

Співвідношення 80:20 для поділу даних є загальноприйнятим стандартом у задачах машинного навчання. Воно забезпечує баланс між обсягом даних для навчання моделі та обсягом даних для тестування, дозволяючи отримати надійні результати оцінки продуктивності моделі.

Компроміс між навчанням і тестуванням:

- Якщо для тестування виділити менше ніж 20% даних, результати можуть бути ненадійними через недостатній розмір тестового набору. Це особливо важливо для задач із високою варіативністю даних, де потрібна репрезентативна вибірка для точного оцінювання продуктивності.
- У випадку, якщо для тестування виділити більше ніж 20%, модель може втратити доступ до значного обсягу даних для навчання. Це може знизити якість навчання, особливо для моделей, які потребують великих наборів даних, таких як LSTM чи Transformers. Як наслідок, зменшується здатність моделі до узагальнення на нових даних.

Таким чином, співвідношення 80:20 дозволяє забезпечити ефективне навчання моделей і водночас отримати надійну оцінку їхньої продуктивності, що робить цей підхід практичним і універсальним для багатьох задач.

Модель LSTM була обрана для задачі прогнозування параметра RESP завдяки її здатності ефективно обробляти часові ряди, зберігаючи інформацію про довгострокові залежності між даними. Архітектура моделі була оптимізована для даної задачі та включала кілька шарів LSTM із налаштованою кількістю одиниць у кожному шарі. Для зменшення ризику перенавчання використовувався Dropout, що забезпечує регуляризацию та підвищує стійкість моделі. Модель 1D-CNN використовувала фіксовані входні послідовності з визначеним розміром вікна, що дозволило застосувати згорткові шари для автоматичного вилучення ключових ознак із сигналів. Такий підхід сприяв виділенню локальних залежностей у даних, забезпечуючи ефективну обробку без потреби у значній кількості попередньої обробки. Архітектура Transformers включала багатоголові механізми уваги, що дозволило моделі враховувати взаємозв'язки між усіма елементами послідовності одночасно. Однак специфіка задачі, орієнтованої на часові ряди з довготривалими залежностями, обмежила продуктивність цієї архітектури. Модель потребувала значних обчислювальних ресурсів і не продемонструвала конкурентної ефективності порівняно з LSTM та 1D-CNN.

Для кожної моделі було проведено ретельну оптимізацію гіперпараметрів з метою покращення її продуктивності та адаптації до специфіки задачі прогнозування параметра RESP. Процес налаштування включав пошук оптимальних значень таких параметрів, як кількість одиниць у прихованих шарах, коефіцієнти Dropout для регуляризації, розміри вікон для моделі 1D-CNN, розміри пакету (batch size) та швидкість навчання (learning rate). Для систематичного перебору параметрів використовувалася стратегія ґраткового пошуку, яка дозволяє дослідити всі можливі комбінації гіперпараметрів у заданому діапазоні. У деяких випадках для зменшення обчислювальних витрат і пришвидшення процесу оптимізації застосовувалася стратегія випадкового пошуку (англ. Random Search),

яка вибирає випадкові комбінації параметрів, забезпечуючи при цьому пошук у широкому просторі можливих значень.

Результати налаштування оцінювалися на валідаційному наборі даних для кожної моделі. Остаточні значення гіперпараметрів обиралися відповідно до найкращої продуктивності за метриками, такими як MSE, MAE та R^2 , що забезпечило надійну узгодженість між навчальними і тестовими наборами.

Зокрема, для LSTM були перевірені різні значення кількості шарів та одиниць у шарі (50, 100, 200) у поєднанні з різними значеннями Dropout (0.2, 0.3, 0.5). Для 1D-CNN тестувалися різні розміри ядра згортки (3, 5) і кількість фільтрів у шарах (32, 64, 128). У Transformers оптимізація включала налаштування кількості голів уваги та розмірності прихованих шарів. Остаточні налаштування були визначені за метриками продуктивності.

4.4. Відбір найкращої моделі машинного навчання для діагностики апное сну

Оцінювання моделей проводилося на тестовому наборі даних за допомогою метрик точності, повноти та F1-міри. Ці метрики є ключовими для задач класифікації, а їх використання в задачі регресії потребувало додаткового перетворення цільової змінної у класи відповідно до певних порогів. Точність (accuracy) вимірювала частку коректно класифікованих зразків серед усіх передбачених. Формально, точність визначається як:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

де TP, TN, FP, FN — кількість істинно позитивних, істинно негативних, хибно позитивних та хибно негативних передбачень відповідно.

Повнота (recall) відображала здатність моделі виявляти всі позитивні випадки. Формула для обчислення повноти має вигляд:

$$Recall = \frac{TP}{TP + FN}$$

F1-міра, як гармонічне середнє точності та повноти, дозволила оцінити баланс між цими двома характеристиками:

$$F1_{score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Додано примітку [6]: Навести в роботі матриці невідповідності (confusion matrix) для класифікації

Розрахувати чутливість та специфічність - під матрицями невідповідності

Додано примітку [7]: де наведено матриці невідповідності та таблиці з показниками чутливості та специфічності діагностики? По тексту досі не зрозуміло, в чому полягала задача класифікації. Норма-апное? На якому часовому проміжку? На кількох прикладах кожного класу навчалась нейронна мережа?

Модель LSTM демонструвала найкращі результати за всіма метриками, досягаючи високої точності, повноти та F1-міри. Це свідчить про здатність моделі забезпечувати баланс між виявленням істинно позитивних значень і мінімізацією хибних передбачень. Модель 1D-CNN також продемонструвала задовільні результати, але поступилася LSTM, зокрема через більшу кількість хибно негативних передбачень. Модель Transformers виявилася найменш ефективною, демонструючи низькі значення точності та повноти, що вказує на обмежену здатність цієї архітектури до роботи з подібними даними.

Навчання моделей на підготовлених даних із подальшою оптимізацією гіперпараметрів дало змогу ідентифікувати найефективнішу архітектуру для прогнозування параметра RESP. Результати оцінки моделей за метриками точності, повноти та F1-міри підтвердили, що LSTM є найкращим вибором для цієї задачі. Ця модель продемонструвала найвищу точність та здатність до узагальнення, забезпечуючи надійне прогнозування навіть на тестових даних. Модель 1D-CNN, хоча й поступається LSTM за продуктивністю, може бути розглянута як прийнятний альтернативний варіант, особливо для сценаріїв, де потрібно зменшити обчислювальні витрати. Водночас модель Transformers, попри сучасність її архітектури та потенціал у багатьох задачах, не продемонструвала достатньої ефективності у прогнозуванні параметра RESP. Це може бути обумовлено особливостями даних, орієнтованих на часові ряди, або недостатньою адаптацією трансформерів до специфіки задачі. Таким чином, результати підтверджують перевагу LSTM як найбільш придатної архітектури для аналізу фізіологічних часових рядів у контексті даної задачі.

4.5. Інтеграція моделей в мікросервісну архітектуру, забезпечення можливості їх оновлення та перенавчання

У представленій архітектурі моделі машинного навчання інтегровані в рамках мікросервісної структури, що забезпечує високу модульність і гнучкість, а також дозволяє легко оновлювати та перенавчати моделі без впливу на інші компоненти системи. Архітектура базується на принципах подієво-орієнтованого підходу (EDA), де кожен мікросервіс має чітко визначені завдання та взаємодіє з

іншими елементами через подієву шину. Це дозволяє ефективно розподіляти обчислювальні ресурси та значно підвищує здатність системи до масштабування.

На початковому етапі дані збираються на рівні edge computing, де пристрій ESP32 виконує первинну обробку біометричних сигналів, зокрема PPG, HR та SpO₂. Зібрані дані передаються до хмарного рівня через мобільний додаток за допомогою BLE-з'єднання. У хмарній інфраструктурі API Gateway виконує роль централізованої точки входу, забезпечуючи маршрутизацію запитів між мікросервісами. Такий підхід дозволяє чітко розмежувати інтерфейс взаємодії з користувачем і внутрішню логіку роботи системи, підвищуючи її гнучкість і масштабованість.

На рівні хмари функціональність системи розподілена між декількома мікросервісами, кожен із яких виконує специфічну роль. Мікросервіс обробки даних займається підготовкою інформації до моделювання, включаючи нормалізацію, фільтрацію та створення часових вікон. Цей компонент забезпечує якісну попередню обробку, після чого дані передаються до мікросервісу зберігання, який використовує базу даних PostgreSQL для збереження як необроблених, так і підготовлених наборів. Одночасно підготовлені дані направляються до мікросервісу штучного інтелекту, який виконує прогнозування, використовуючи інтегровану модель машинного навчання. Такий розподіл забезпечує модульність і високу ефективність системи.

Мікросервіс штучного інтелекту виконує ключову роль у системі, реалізуючи обчислювальні функції для прогнозування на основі підготовлених даних. Він розгортається як автономний модуль, що забезпечує можливість оновлення або перенавчання моделі без впливу на інші компоненти системи. Отримані результати передаються до мікросервісу обробки результатів, де здійснюється їх пост-обробка, після чого дані зберігаються у сервісі зберігання результатів. Це дозволяє використовувати прогнозовані дані для подальшого аналізу, візуалізації або інтеграції з іншими компонентами системи, забезпечуючи гнучкість і функціональність архітектури.

В архітектурі особливий акцент зроблено на можливості оновлення та перенавчання моделей штучного інтелекту. Дані, які постійно акумулюються у базі даних, слугують основою для перенавчання моделей, що дозволяє адаптувати їх до нових умов або збільшення обсягу даних. Для цього передбачено інтеграцію автоматизованих процесів перенавчання, які активуються за допомогою спеціальних подій у подієвій шині. Наприклад, після досягнення певного порогу нових даних система може ініціювати перенавчання моделі та її автоматичне повторне розгортання у відповідному мікросервісі. Такий підхід забезпечує безперервну оптимізацію моделі, підтримуючи високу продуктивність і точність прогнозів без зупинки роботи всієї системи.

Архітектура включає інтегровані механізми моніторингу точності моделей, які сприяють підтримці їхньої продуктивності на високому рівні. У рамках сервісу обробки результатів регулярно обчислюються метрики ефективності, такі як Mean Squared Error (MSE), Mean Absolute Error (MAE) та коефіцієнт детермінації (R^2). Ці показники дозволяють оцінювати продуктивність моделі після кожного оновлення або перенавчання. У разі виявлення зниження точності система може автоматично ініціювати процес перенавчання моделі або надсилати сигнал розробнику для подальшого аналізу та втручання. Такий підхід забезпечує не лише високу якість прогнозів, але й оперативну адаптацію моделі до змін у даних.

Основними перевагами запропонованої архітектури є її модульність і масштабованість, що забезпечують високу гнучкість у роботі системи. Завдяки розподіленню функціональності між мікросервісами, оновлення окремих компонентів може виконуватися незалежно, без впливу на загальну роботу системи. Інтеграція моделей машинного навчання як окремих мікросервісів відкриває можливість одночасного використання різних алгоритмів, що дозволяє обрати найефективніше рішення для кожної конкретної задачі. Такий підхід сприяє адаптації системи до змін вимог, розширенню її функціональності та впровадженню нових технологій, зберігаючи при цьому її продуктивність та стабільність.

РОЗДІЛ 5. РОЗРОБКА НОСИМОГО ПРИСТРОЮ ДЛЯ РЕЄСТРАЦІЇ ТА ЗБОРУ ФПГ СИГНАЛУ

Розроблений носимий пристрій представляє собою сучасне рішення для безперервного моніторингу біометричних параметрів пацієнтів у реальному часі. Завдяки використанню мікроконтролера ESP32-C3 та сенсора Max30102 пристрій забезпечує високу точність вимірювань і стабільність роботи навіть у складних умовах. Інтеграція з енергоефективними технологіями дозволяє пристрою працювати автономно протягом тривалого часу, зберігаючи компактність і низьке енергоспоживання. Додатковою перевагою є інтеграція з хмарними сервісами, що забезпечує можливість обробки, аналізу та зберігання зібраних даних для медичного використання, що значно розширює функціональні можливості пристрою.

Центральним елементом пристрою виступає мікроконтролер ESP32-C3, який відзначається низкою переваг, зокрема 4 МБ флеш-пам'яті, вбудованими модулями WiFi та Bluetooth, а також високою енергоефективністю. Ці характеристики забезпечують можливість локального збереження даних у випадку відсутності підключення до мережі, передачу інформації на сервер через бездротові протоколи зв'язку, а також загальну зручність у використанні. Автономна робота пристрою реалізується за допомогою літій-іонної батареї ємністю 100 мА·год, зарядка якої здійснюється через модуль TP4056. Цей модуль забезпечує захист батареї від перезарядження та надмірного розрядження, гарантуючи стабільність та надійність роботи пристрою навіть у тривалих циклах експлуатації.

Сенсор Max30102 використовується для вимірювання ключових біометричних параметрів, включаючи частоту серцевих скорочень (HR), рівень насичення крові киснем (SpO_2) та фотоплетизмограму (PPG). Проте експериментальні дослідження показали, що отримані сирі дані часто характеризуються високим рівнем шумів, спричинених зовнішніми чинниками, такими як піт, рухи пацієнта або вплив зовнішнього освітлення. Для усунення цих проблем була розроблена система фільтрації сигналів, основним компонентом якої

є фільтр Кальмана. Завдяки цьому вдалося значно підвищити якість сигналів, ефективно зменшивши вплив шумів. Такий підхід забезпечив високу точність вимірювань, що є критично важливим для застосування пристрою в медичних системах.

Особливий акцент у розробці пристрою зроблено на оптимізації енергоспоживання. Було реалізовано циклічний режим роботи, за якого пристрій активно записує дані протягом 2 хвилин, після чого переходить у сплячий режим на 20 хвилин. Такий підхід дозволив значно знизити енергоспоживання, забезпечуючи до 48 годин автономної роботи на одному заряді батареї. Крім того, пристрій оснащено спеціальним тригером, який автоматично активує запис даних поза основним циклом у разі різких змін інтенсивності інфрачервоного сигналу. Це рішення дає змогу реєструвати критичні події, такі як активні рухи пацієнта, і забезпечує більш детальний і точний моніторинг фізіологічного стану.

Апаратна схема пристрою спроектована на основі взаємодії ключових компонентів, таких як мікроконтролер ESP32-C3, сенсор Max30102, зарядний модуль TP4056 і стабілізатор напруги. Живлення забезпечується літій-іонною батареєю ємністю 100 мА·год. Для стабільної роботи всіх компонентів напруга батареї стабілізується до 3.3 В перед подачею на ESP32-C3 і Max30102. Сенсор Max30102 підключений до ESP32 через стандартний інтерфейс I²C, використовуючи GPIO4 (SCL) і GPIO5 (SDA) для передачі даних і синхронізації. Всі компоненти з'єднані спільним контуром заземлення, що гарантує надійну та стабільну роботу пристрою, мінімізуючи можливі перешкоди в роботі електронних компонентів.

На рис. 5.1. зображено схемотехнічний дизайн пристрою, що відображає фізичну інтеграцію ключових компонентів: мікроконтролера ESP32-C3, сенсора Max30102, зарядного модуля TP4056 та літій-іонної батареї. Дизайн включає всі необхідні з'єднання, такі як інтерфейс I²C між ESP32-C3 та Max30102, стабілізацію напруги за допомогою регулятора, а також підключення батареї до зарядного модуля. Представлена схема деталізує електричну архітектуру пристрою, забезпечуючи надійне функціонування та взаємодію між компонентами.

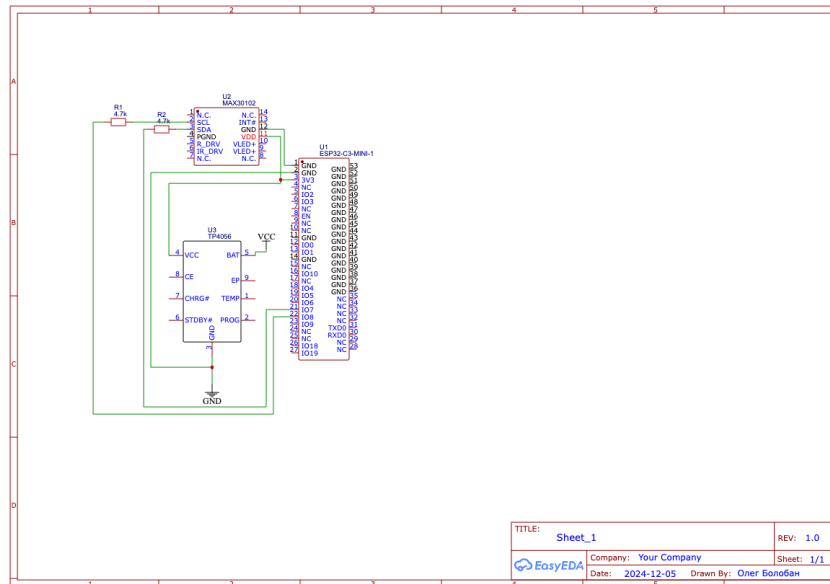


Рис. 5.1. – Схемотехнічний дизайн пристрою

Результати роботи пристрою підтверджують його ефективність у реальних умовах експлуатації. Зібрані та попередньо оброблені дані успішно передаються на мобільний додаток або безпосередньо до хмарної платформи через WiFi чи Bluetooth для подальшого аналізу. У випадках тимчасової відсутності мережевого з'єднання пристрій зберігає дані у вбудованій флеш-пам'яті мікроконтролера ESP32-C3, забезпечуючи їх доступність для синхронізації після відновлення підключення. Такий підхід гарантує безперервність збору та зберігання інформації, навіть за умов нестабільного зв'язку.

Таким чином, створений пристрій ефективно поєднує високу функціональність, точність вимірювань та енергоефективність. Його компактний розмір і модульна архітектура ідеально підходять для інтеграції в носимі системи моніторингу здоров'я, забезпечуючи тривалий автономний контроль стану пацієнтів у режимі реального часу. Такий підхід відкриває нові можливості для

розвитку персоналізованої медицини та дистанційного моніторингу, що є ключовими елементами сучасної охорони здоров'я.

У розробленій системі носимий пристрій використовує внутрішню флеш-пам'ять об'ємом 4 МБ, вбудовану в мікроконтролер ESP32-C3, для тимчасового зберігання даних. Ця пам'ять служить буфером для зібраних біометричних показників, таких як частота серцевих скорочень, рівень насичення крові киснем і фотоплетизмограма. У випадку втрати підключення до мобільного пристрою дані записуються у вигляді компактних бінарних файлів, що забезпечує оптимальне використання доступного обсягу пам'яті. Завдяки використанню динамічного буфера реалізується механізм циклічного перезапису, який запобігає переповненню пам'яті та гарантує збереження лише актуальних даних, готових для передачі після відновлення зв'язку.

5.1. Безпека передачі даних від носимого пристрою до мобільного

Передача зібраної інформації з носимого пристрою на мобільний додаток здійснюється через Bluetooth Low Energy (BLE). Процес передачі включає кілька етапів, починаючи з аутентифікації мобільного пристрою. Для цього під час першого підключення генерується унікальний криптографічний ключ, який зберігається у внутрішній пам'яті ESP32-C3. Цей ключ передається на мобільний додаток через QR-код, який сканується користувачем. QR-код містить криптографічний ключ та унікальний ідентифікатор пристрою, що забезпечує надійність процесу первинної аутентифікації.

Дані передаються у зашифрованому вигляді за допомогою алгоритму XOR-шифрування, який гарантує високий рівень безпеки від перехоплення або несанкціонованого доступу. Шифрування виконується на стороні носимого пристрою перед передачею через BLE. Дані передаються у вигляді невеликих пакетів, кожен з яких містить контрольну суму для перевірки цілісності на стороні мобільного додатку. У разі пошкодження або втрати пакета передбачено механізм повторної передачі, що мінімізує ймовірність втрати даних.

На мобільному пристрої отримані дані декодуються за допомогою криптографічного ключа, отриманого з QR-коду. Після декодування дані

зберігаються у внутрішній базі мобільного додатку або передаються у хмарний сервіс для подальшого аналізу. У процесі декодування використовуються додаткові перевірки для забезпечення відповідності даних очікуваному формату.

XOR-шифрування полягає в побітовій операції XOR між байтами даних і ключа. Один і той самий ключ використовується як для шифрування, так і для розшифрування.

$$result = data \oplus key$$

XOR-шифрування є одним із найпростіших методів захисту даних, який широко використовується в системах з обмеженими обчислювальними ресурсами. Цей метод заснований на побітовій операції виключного OR (XOR) між кожним байтом даних і ключа. Основною особливістю XOR-шифрування є його симетричність, що дозволяє використовувати той самий ключ як для шифрування, так і для розшифрування. Завдяки цьому алгоритм надзвичайно простий в реалізації, що робить його придатним для використання в мікроконтролерах, таких як ESP32, які обмежені в обчислювальних потужностях та пам'яті.

Однією з ключових переваг XOR-шифрування є його простота реалізації, яка дозволяє легко інтегрувати цей метод у пристрої з обмеженими ресурсами. Мінімальні вимоги до пам'яті та процесора роблять XOR-шифрування ідеальним для мікроконтролерів, таких як ESP32. Швидкість роботи алгоритму забезпечує низьку затримку під час шифрування або розшифрування, що є важливим для реального часу.

Додатково, симетричний характер алгоритму дозволяє уникнути складної взаємодії між шифрувальними і розшифрувальними системами. Передача ключа через QR-код забезпечує додатковий рівень безпеки, оскільки ключ не передається через відкриті канали зв'язку. Завдяки цим перевагам XOR-шифрування може бути використане у багатьох задачах із базовими вимогами до безпеки.

Процес XOR-шифрування працює швидко, оскільки не потребує складних математичних операцій. Крім того, метод не вимагає додаткових бібліотек чи складної конфігурації, що значно спрощує його інтеграцію у вбудовані системи. Для забезпечення базового рівня захисту ключ шифрування зберігається в пам'яті

пристрою, а для аутентифікації користувача та передачі ключа використовується QR-код, що забезпечує зручність використання.

З метою забезпечення високого рівня безпеки інформації ключ кодування зберігається у захищеній пам'яті ESP32-C3, доступ до якої неможливий через зовнішні інтерфейси. Всі операції, пов'язані з обміном даними, здійснюються лише з авторизованими пристроями, які пройшли первинну аутентифікацію. Захист від перехоплення даних забезпечується шифруванням у каналі передачі. Крім того, моніторинг метаданих про підключення дозволяє виявляти та блокувати підозрілі спроби доступу.

У результаті реалізовані механізми забезпечують безпечну передачу чутливих даних, високу точність зберігання інформації та зручність користування. Локальне зберігання даних у пам'яті пристрою забезпечує безперервність моніторингу навіть за відсутності підключення до мобільного пристрою. Шифрування даних та використання криптографічних ключів гарантує, що дані будуть доступні лише авторизованим пристроям, що робить систему надійною та безпечною для користувача. Такий підхід забезпечує надійну платформу для збору, обробки та передачі біометричних даних, яка може бути інтегрована у різноманітні системи моніторингу здоров'я.

5.2. Оптимізація програмного забезпечення для забезпечення енергоефективності та скорочення часу виконання

Щоб система на ESP32-C3 працювала достатньо швидко та водночас споживала мінімум енергії, потрібно одночасно опрацьовувати три рівні — апаратний, системний і алгоритмічний. Платформа ESP32-C3 підтримує чотири стани живлення (Active → Modem-sleep → Light-sleep → Deep-sleep). Основний сценарій — працювати в активному стані лише під час зчитування даних з сенсорів й обробки сегмента, а решту часу перебувати в режимі легкого сну (падіння споживання з $\approx 80\text{mA}$ до $< 1\text{mA}$). Для пауз довші за 5с варто переходити в режим глибокого сну ($\approx 20\mu\text{A}$) з пробудженням по таймеру або GPIO-перериванню. Живлення Wi-Fi та Bluetooth слід вимикати за допомогою методів `esp_wifi_stop()` або `esp_bt_controller_disable()` у ті цикли, де передача даних не відбувається.

На системний рівні за допомогою бібліотеки `esp-idf` активується динамічне зниження тактової частоти (`CONFIG_PM_ENABLE`), щоб ядро автоматично уповільнювалось з 160 МГц до 20 МГц, коли задача чекає подій. Код, що виконується в обчислювальних піках (фільтр Калмана, детектор піків), розміщено у швидкій внутрішній пам'яті (`IRAM_ATTR`) і компілюється з прапорцем `-Ofast`; рідко викликані функції записані у швидкій пам'яті, щоб не збільшувати споживання SRAM. Для I²C-обміну з MAX30102 виставлено частоту 100 кГц — вища частота майже не скорочує час сесії, зате підвищує струм на шині.

На алгоритмічному рівні енергоефективність забезпечується за допомогою зменшення тактових циклів. Замість плаваючої арифметики використовується Q15-фіксовану; коефіцієнти Калман-фільтра попередньо обчислені на іпрошиті константами. На мікроконтролері виконується лише фільтрація й первинний детектор піків; це скорочує час активного використання радіомодуля приблизно втричі. Сам датчик налаштований на 100sps, увімкнено в імпульсний режим за вимогою і знижено струм світлодіода до межі 7 мА, щоб споживати мінімум електроенергії, але достатньо, щоб зчитувати сатурацію.

Споживання електроенергії у режимі «нічного моніторингу» триває 100 мс в активному режимі кожні 2 с, що становить 5% робочого циклу, інші 95% часу використовується режим легкого сну з силою струму < 1 мА. Сумарно ≈ 3 мА середнього струму; на акумуляторі 500 мА·год — понад 6 діб автономної роботи.

Таким чином, поєднання глибокого сну, динамічного зниження частоти, мінімального радіоефіру та фіксованої арифметики забезпечує оптимізацію програмного забезпечення, енергоефективність, скорочення часу виконання й обчислення з низьким енергоспоживанням — саме те, що потрібно для тривалої роботи системи з живленням від батареї.

5.3. Формальна специфікація та перевірка часової логіки системи засобами часової логіки дій

Щоби не покладатися лише на експериментальні тести, для критичних систем рекомендовано формально описати очікувану поведінку та автоматично перевірити, чи не з'являться помилки за будь-якої послідовності подій. У цій роботі така перевірка виконана мовою часової логіки дій (англ. Temporal Logic of Actions (TLA)), а коректність специфікацій перевірена перевірником моделей (англ. TLA Checker (TLC)).

У формальній специфікації TLA+ описано повний ланцюг обробки медичних даних та класифікації апное: «сенсор – мікроконтролер – смартфон – хмара» як єдину послідовність дій (табл. 5.1.): початкова конфігурація фіксує пустий буфер датчика, вимкнений радіомодуль і достатній заряд акумулятора, тоді як наступні переходи задають дозволені кроки—зчитування імпульсу з фотоплетизмографа, копіювання відліків у пам'ять ESP32-C3, короткочасне ввімкнення BLE для передавання сегмента, локальну класифікацію на смартфоні та запис результату у хмарну базу. Для кожного етапу темпоральні формули задають інваріанти безпеки («дані ніколи не губляться», «радіоканал не вмикається без зашифрованого пакета») і вимоги властивості прогресу («кожен 30-секундний сегмент отримує мітку менш як за секунду, а push-сповіщення надходить користувачеві менш як за п'ять секунд»). TLC перебирає всі можливі порядки подій, підтверджуючи, що за будь-якого сценарію не виникає переповнення буфера, не порушуються часові межі та батарея не розряджається до завершення сесії, отже, задана часово-логічна поведінка системи є коректною.

Таблиця 5.1 Опис рівнів системи класифікації апное для оцінки TLA

| Рівень | Приклад станів | Ключові вимоги |
|-----------------|-------------------|---|
| Датчик MAX30102 | FIFO, LED_State | Безпека: FIFO ніколи не переповнюється. Прогрес: кожен імпульс рано чи пізно зчитує ESP32-C3. |
| ESP32-C3 | Buf, BLE, Battery | Безпека: передається лише зашифрований пакет. Прогрес: |

| | | |
|----------|--------------|---|
| | | сегмент довжиною 30 с потрапляє у смартфон не пізніше ніж за 35 с. |
| Смартфон | Queue, Model | Безпека: класифікація запускається лише над валідним сегментом. Прогрес: кожен сегмент отримує мітку ≤ 1 с. |
| Хмара | RabbitMQ, DB | Безпека: жодне повідомлення не губиться. Прогрес: результат доходить до користувача ≤ 5 с. |

У формальній моделі часової логіки дій початковий стан («Init») фіксує значення всіх змінних: буфер датчика порожній, радіомодуль BLE вимкнено, а рівень заряду акумулятора не нижчий за 450 мА·год. Подальшу еволюцію системи задає оператор «Next», який описує допустимі переходи; зокрема, перехід BLE_On допускається лише тоді, коли у буфері вже накопичилися дані й радіоканал досі вимкнений, після чого BLE переводиться у стан On, а заряд батареї зменшується на одну одиницю. Для такої системи сформульовано два види вимог. Інваріант безпеки гарантує, що дані або містяться в буфері, або передаються каналом BLE, тобто буфер ніколи не залишається непорожнім, коли радіо неактивне. Властивість прогресу фіксує, що кожен новий 30-секундний фрагмент сигналу обов'язково отримує класифікаційну мітку менш як за секунду. Перевірник моделей TLC проаналізував понад два мільйони комбінацій станів і не виявив контрприкладів, отже, сформульовані часові й безпекові обмеження виконуються за будь-якої послідовності подій. Такий результат доповнює експериментальні випробування й математично доводить: буфер ніколи не переповнюється і дані не губляться навіть за пікових навантажень; акумулятор не розряджається до критичного рівня раніше, ніж сегмент буде передано й оброблено; усі затримки залишаються в межах однієї секунди для класифікації та п'яти секунд для push-сповіщення. Формальна специфікація TLA+ разом із перевіркою TLC таким чином забезпечує строгі гарантії коректності часової логіки й енергоспоживання системи, не вимагаючи детального аналізу кожної гілки вихідного коду.

Перевірено TLA-модуль, у якому формалізовано роботу підсистеми «датчик MAX30102 → ESP32-C3 → смартфон» засобами TLC охопила 4515512 згенерованих конфігурацій, з-поміж яких 1006008 виявилися унікальними; у черзі

залишилося лише 2004 недосліджених стани, що свідчить про майже повне покриття простору. Максимальна довжина безповторного шляху в графі досягла 1003 переходів, середній вихідний степінь вершини становив 1 (зафіксовано не більш як три альтернативні дії, причому у 95 % випадків — лише одну), отже поведінка системи близька до детермінованої. Увесь аналіз було виконано за 40 с; жодного контрприкладу TLC не виявив, тож як інваріант безпеки «NoLostData», так і властивість прогресу «TimelyClass» залишаються істинними на всьому дослідженому просторі станів. Практично це означає, що буфер жодного разу не виявився непорожнім за вимкненого BLE-каналу, а кожен новий сегмент сигналу незмінно діставав класифікаційну мітку менш ніж за секунду, незалежно від порядку внутрішніх подій; отже часово-логічна коректність і відсутність втрати даних формально доведені.

РОЗДІЛ 6. РЕАЛІЗАЦІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ДІАГНОСТИЧНИХ РІШЕНЬ ДЛЯ ВІДКРИТТЯ АПНОЕ СНУ

6.1. Розробка енергоефективної прошивки ESP32-C3 для зчитування даних з ФПГ сенсору, фільтрація сигналів і шифрованої BLE-передачі сегментів.

Для отримання клінічно придатного ФПГ сигналу, прошивка мікроконтролера (Додаток А) повинна забезпечити рівномірне зчитування даних з ФПГ сенсору з інтервалом 10мс, водночас не виходячи за суворі енергетичні рамки носимого пристрою. З цією метою основний цикл роботи зчитування даних мікроконтролером ESP32-C3 організовано як об'єднання коротких активних фаз ($\approx 0,13$ мс на читання FIFO та фільтрацію) і довгих проміжків Light-sleep, а кожне переривання сенсора слугує єдиним тригером для всіх подальших обчислень. Переривання (англ. interrupt (INT)) - це спеціальний вивід-сигнал, який переходить у активний стан щоразу, коли в модулі виникає подія, визначена у його регістрі переривань. Найчастіше такою подією є заповнення чергового відліку у FIFO-буфері. Підключивши вивід до вводу-переривання мікроконтролера, девайс може

виходити зі сну лише тоді, коли дійсно надійшли нові дані, — це усуває потребу в постійному опитуванні шини I²C і суттєво економить енергію.

У програмному забезпеченні носимого пристрою запит частоти 100 Гц реалізовано як подію-керований цикл, у якому апаратне переривання INT датчика ФПГ сенсору є єдиним тригером для всіх подальших дій. Після ініціалізації INT-вивід під'єднується до цифрового виходу GPIO 3, а обробник переривання обмежується встановленням прапорця `irqFlag`, що мінімізує час перебування ядра в контексті обробника переривань (англ. Interrupt Service Routine (ISR)) і повертає його у режим Light-sleep протягом декількох мілісекунд.

У головному циклі роботи контролер перевіряє, чи минув черговий інтервал у 10мс і чи піднято прапорець `irqFlag`. Якщо обидві умови виконано, з FIFO сенсора по I²C (400 кГц) зчитується черговий 18-розрядний відлік ІЧ-каналу; операція триває $\approx 120\mu\text{s}$. Зразу після зчитування сигнал піддається онлайн-фільтрації одноканальним Калман-фільтром, реалізованим у Q15-арифметиці: коефіцієнт Калмана обчислюється на цілих числах, а множення та нормування здійснюються побітовими зсувами, що зменшує час обробки одного зразка до $\approx 8\mu\text{s}$ без залучення апаратного співпроцесора для десяткових дробів (англ. Floating-Point Unit (FPU)). Значення після фільтрації заноситься до буфера сегмента; коли лічильник досягає 300 зразків (30с), сегмент вважається завершеним.

По завершенні вікна виконуються два варіанти дій, залежно від статистичних характеристик: якщо дисперсія перевищує заданий поріг, сегмент маркується як «аномальний» і у BLE-пакет вноситься відповідний тег; інакше передається лише усереднена телеметрія. Для передачі даних використовується характеристика «Вимірювання частоти серцебиття». Первісні п'ять байтів містять службові прапорці та значення частоти серцебиття і насичення крові киснем у форматі Q15, наступні шістнадцять байтів — зразки інфрачервоного сигналу, після чого весь пакет шифрується за схемою AES-128 CCM, а в кінець додається чотирибайтовий тег для перевірки цілісності. Відправлення такого захищеного пакета через функцію `esp_ble_gatts_send_indicate` займає приблизно 4 мс радіочасу.

Щоб досягти високої енергоефективності, ядро активно лише 1,3 % кожного 10-мс слоту (читання, фільтрація, буферизація); решту часу воно перебуває у Light-sleep, де тактова частота автоматично падає до 20 МГц, а споживаний струм — до ≈ 1 мА. Після відправлення 30-секундного сегмента контролер занурюється у Deep-sleep на чотири хвилини, залишаючи живими лише RTC-таймер і лінію INT; струм у цьому режимі не перевищує 20 μ А. Такий розподіл— $\approx 0,13$ мс активних обчислень, $\approx 9,9$ мс Light-sleep та тривалі блоки Deep-sleep—забезпечує середній робочий струм ≈ 3 мА, завдяки чому пристрій з акумулятором 500 мА·год працює понад шість діб без підзаряджання й при цьому підтримує гарантований інтервал вибірки $10 \pm 0,01$ мс, необхідний для коректної реконструкції ФПГ-сигналу та наступної хмарної класифікації апное.

У функції `setup()` перед початком реклами виконується читання 48-бітної MAC-адреси центрального контролера, яка потім перетворюється в десятковий рядок і додається до префікса “НМ_”. Наприклад, MAC-адреса 0x0F A1 B2 C3 D4 E5 перетворюється на рядок НМ_264938245159012. Це ім'я передається стеку Bluetooth Low Energy перед початком реклами. Далі налаштовується пакет рекламних даних із включенням імені та UUID сервісу «Вимірювання частоти серцебиття», після чого викликається `esp_ble_gap_start_advertising()`.

У зворотному виклику обробника `gap_event_handler` система отримує підтвердження успішної конфігурації рекламних даних і старту реклами. Поки не встановлено з'єднання, у циклі `loop()` мікроконтролер здійснює мигання світлодіода, очікуючи переривання `ESP_GATTS_CONNECT_EVT`. Після підключення мобільного пристрою змінюються параметрами з'єднання для оптимального співвідношення швидкості та енергоспоживання.

Якщо ж після надсилання сегмента система переходить у глибокий сон, перед сном поточний стан з'єднання (MAC-адреса центрального пристрою та індекс транзакції) зберігається у флеш-пам'яті ESP32-C3. Після пробудження за допомогою RTC-таймера прошивка завантажує збережені дані та в автоматичному режимі намагається відновити Bluetooth-з'єднання з тим самим мобільним

пристроєм — це забезпечує мінімальну затримку на повторну ініціалізацію реклами та повторне підключення без участі користувача.

```
// Зчитування MAC-адреси та формування імені пристрою у десятковому форматі
uint64_t getChipId() {
    uint64_t mac = ESP.getEfuseMac(); // 48-бітна MAC-адреса
    uint64_t id = 0;
    for (int i = 0; i < 6; ++i) {
        // Зсув байтів у зворотному порядку та нарощування у 48-бітне число
        id |= ((mac >> (40 - i * 8)) & 0xFF) << (i * 8);
    }
    return id; // Наприклад, 264938245159012
}

void setup() {
    Serial.begin(115200);
    delay(1000);

    // Отримуємо унікальний ідентифікатор і перетворюємо його на рядок
    uint64_t chipId = getChipId();
    char deviceName[32];
    // Формат "HM_<decimal_id>"
    snprintf(deviceName, sizeof(deviceName),
             "HM_%llu", (unsigned long long)chipId);
    Serial.print("Device name: ");
    Serial.println(deviceName); // HM_264938245159012

    // Далі -- ініціалізація BLE з цим іменем...
    esp_ble_gap_set_device_name(deviceName);
    // ...
}
```

У програмному забезпеченні реалізовано низку додаткових механізмів енерго-ефективності (табл. 6.1.). Для датчика MAX30102 передбачено динамічне керування струмом світлодіодів: під час накопичення 30-секундного сегмента світлодіоди працюють із током 7 мА, тоді як у паузах між зборами їхній струм програмно зменшується до нуля через запис у відповідний регістр сенсора.

Таблиця 6.1. Енергоспоживання мікроконтролера з урахуванням доданих механізмів

| Режим роботи | Тривалість | Струм, мА | Примітка |
|----------------------------------|------------|-----------|-------------------|
| Active (опитування + фільтрація) | 0,13 мс | 90 | TC + Kalman (Q15) |

| | | | |
|---|----------|------|---|
| BLE-передача сегмента (24 байт) | 4 мс | 100 | захищене Notify (AES-CCM) |
| LED-стрім MAX30102 (вікно збирання) | 30 с | 7 | постійний струм світлодіода |
| LED-вимкнення (очікування та сон) | 4...2 хв | 0 | регульований залежно від якості сигналу |
| Light-sleep між ітераціями (10 мс × 3000) | 29,61 с | 1 | CPU 20 МГц, BLE stand-by |
| Deep-sleep після сегмента | 2...4 хв | 0,02 | RTC-таймер + збереження в NVS |
| Перезапуск та відновлення з'єднання | 10 мс | 60 | ініціалізація BLE, відновлення реклами та GAP-подій |

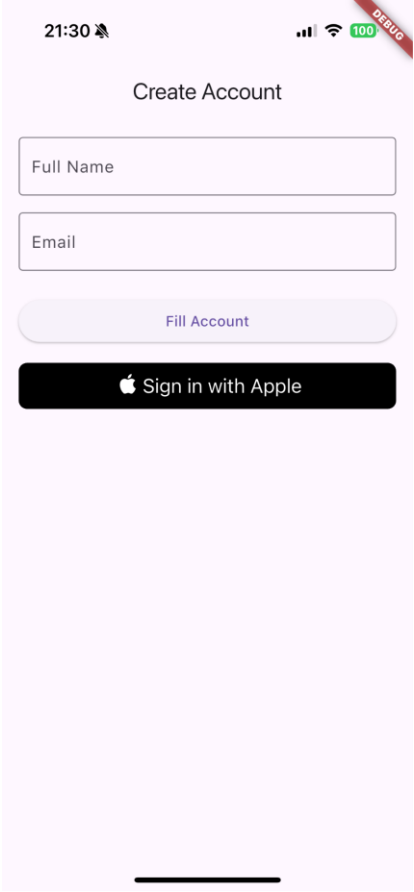
Кожні п'ять хвилин виконується вимір напруги живлення за допомогою функції `readBattery()`, результати якого зберігаються в непохідній пам'яті (NVS) та передаються по захищеному каналу BLE у мобільний застосунок для моніторингу рівня заряду. Також впроваджено адаптивний алгоритм тривалості глибокого сну: якщо в останньому сегменті виявлено менше ніж 10% аномальних відхилень, Deep-sleep триває 240с, а за частіших артефактів—120с, що оптимізує співвідношення енергозбереження та якості даних. Нарешті, перед переходом у Deep-sleep зберігаються ключові змінні (`deviceConnected`, `sampleIndex`, `chipId`) у флеш-пам'яті, що дозволяє після пробудження швидко відновити BLE-з'єднання та настрої сенсора без необхідності повторної ініціалізації.

6.2. Створення мобільного застосунку для приймання BLE-даних, відображення телеметрії та ретрансляції сегментів у хмару.

Для розробки мобільного застосунку було обрано фреймворк Flutter завдяки його здатності єдиним кодом підтримувати і Android, і iOS, що скорочує витрати на розробку та технічну підтримку. У роботі використано пакет `flutter_blue_plus` для взаємодії через Bluetooth Low Energy, який забезпечує надійну фонову синхронізацію, виявлення та автоматичне відновлення з'єднання з носимим

пристроєм після виходу ESP32-C3 із режиму глибокого сну. Для початкового зв'язування застосовано `qr_code_scanner`, що дозволяє сканувати унікальний QR-код пристрою і зберігати його ідентифікатор у захищеному сховищі. Пакет `http` використовується для формування та відправки JSON-запитів до API Gateway, через який сегменти даних маршрутизуються до мікросервісів бекенду. Така комбінація інструментів гарантує безшовну інтеграцію фізичних датчиків із хмарною аналітикою та надає користувачу інтуїтивний і безперервний досвід спостереження за станом здоров'я.

Екран реєстрації реалізовано (рис. 6.1.) у `register_screen.dart` та призначено для введення та валідації даних нового користувача. Під час завантаження віджета за допомогою пакета `provider` створюється екземпляр `UserService`, через який надалі виконуються HTTP-запити на створення облікового запису. Інтерфейс містить поля для імені, адреси електронної пошти, пароля та його підтвердження, кожне з яких перевіряється на потрібну довжину, формат і відповідність через методи `Validators`. Після натискання кнопки «Зареєструватися» відбувається асинхронний виклик `userService.register()`, який повертає токен аутентифікації. У разі успіху токен одразу зберігається в `Flutter Secure Storage`, а застосунок без додаткових кроків переходить на головний екран («дашборд»), що забезпечує плавність користувацького досвіду без необхідності повторного входу. Також для вже зареєстрованих користувачів доступний екран входу і служить для аутентифікації. На цьому екрані користувач вводить електронну пошту та пароль, причому кожне поле контролюється контролером `TextEditingController` і валідується за допомогою тих самих утиліт `Validators`. При натисканні кнопки «Увійти» виконується метод `userService.login()`, який надсилає запит із JSON-тілом до кінцевої точки `/auth/login`. У разі успішного відповіді серверу повернений токен зберігається у захищеному сховищі на мобільному пристрої, а потім відбувається навігація на головний екран. У випадку помилки аутентифікації на екран виводиться сповіщення за допомогою `ScaffoldMessenger`, що інформує користувача про неправильно введені дані, забезпечуючи прозорість і коректне відпрацювання помилкових сценаріїв.



21:30

Create Account

Full Name

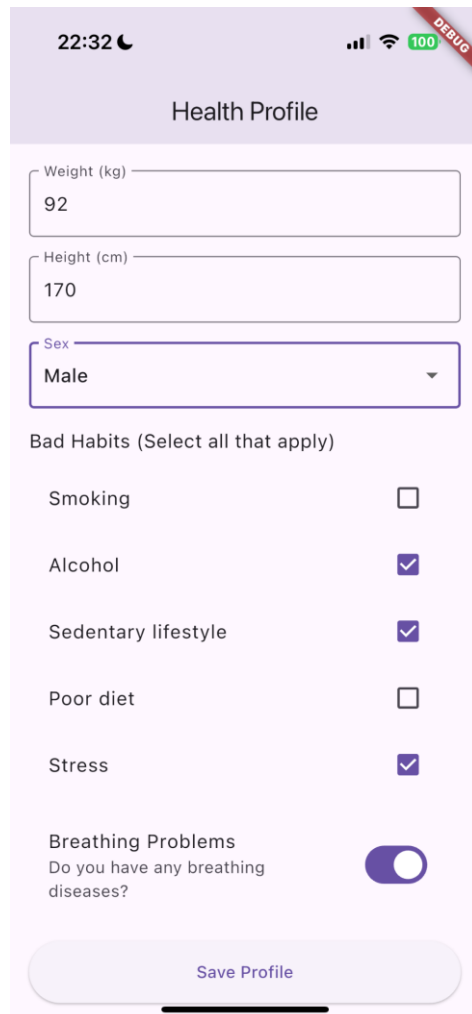
Email

Fill Account

Sign in with Apple

Рис. 6.1. Екран реєстрації користувача

Далі користувач потрапляє на екран налаштування профілю здоров'я (рис. 6.2.), який забезпечує збір ключових персональних даних користувача, необхідних для адаптації алгоритмів класифікації та розрахунку індивідуальних порогів. Інтерфейс складається з полів для дати народження, ваги (кг), зросту (см), статі (чоловік/жінка/інше), переліку шкідливих звичок (множинний вибір) та наявності дихальних проблем (прапорець «так/ні»).



22:32

Health Profile

Weight (kg)
92

Height (cm)
170

Sex
Male

Bad Habits (Select all that apply)

Smoking ☐

Alcohol ☒

Sedentary lifestyle ☒

Poor diet ☐

Stress ☒

Breathing Problems
Do you have any breathing diseases? ☒

Save Profile

Рис. 6.2. Екран налаштування профілю користувача

Для управління станом форми застосовано пакет `provider`: при ініціалізації віджета створюється модель `ProfileViewModel`, яка містить контролери полів і методи валідації. Кожне поле перевіряється на коректність введення (наприклад, вага та зріст — у межах реалістичних фізіологічних значень, дата народження — не пізніше за поточний день).

Після заповнення всіх обов'язкових полів натисканням на кнопку «Зберегти» виконується асинхронний виклик `ApiService.saveProfile(profileData)`, у якому зібрані дані пакуються в JSON-об'єкт. Під час відправки відображається індикатор завантаження, а у разі успішної відповіді сервер повертає статус 200 і оновлений профіль користувача, який зберігається локально за допомогою Flutter Secure Storage. Після збереження застосунок автоматично виконує навігацію до головного екрану (Dashboard), що гарантує цілісність даних і виключає дублювання кроків. Таким чином, Profile Setup Screen виконує важливу функцію первинного налаштування системи: він створює основу демографічних та фізіологічних параметрів, що надалі впливають на точність і персоналізацію прогнозу апное сну. Після першого входу користувач вводить базову інформацію: дату народження, вагу, зріст, стать, шкідливі звички й наявність респіраторних проблем. Ці дані потрібні для корекції алгоритмів класифікації й розрахунку індивідуальних порогів. Успішне збереження через `ApiService` завершує налаштування та повертає на екран сканування QR коду для з'єднання з пристроєм реєстрації ФПГ сигналу.

Екран сканування QR-коду (рис. 6.3.) призначений для швидкого й надійного з'єднання з пристроєм за допомогою унікального ідентифікатора. Після натискання кнопки «Сканувати QR-код» виконується ініціалізація плагіна `qr_code_scanner`, який у реальному часі обробляє відеопотік із камери смартфона та розпізнає коди у форматі QR. Алгоритм роботи складається з трьох етапів. Першим кроком після запуску віджета відбувається налаштування контролера камери з автоматичною фокусуванням і підтримкою лише формату `QR_CODE` для підвищення швидкості та точності сканування. По-друге, при виявленні коду відбувається негайна зупинка потоку та перевірка рядка на відповідність регулярному виразу `"NM_d{12,}$"`, що гарантує коректний формат ідентифікатора `chipId`. По-третє, у разі успіху форматування відсканований номер зберігається через `SecureStorage`, а застосунок автоматично здійснює навігацію до екрану `BleConnectionScreen`, передаючи збережений ідентифікатор.

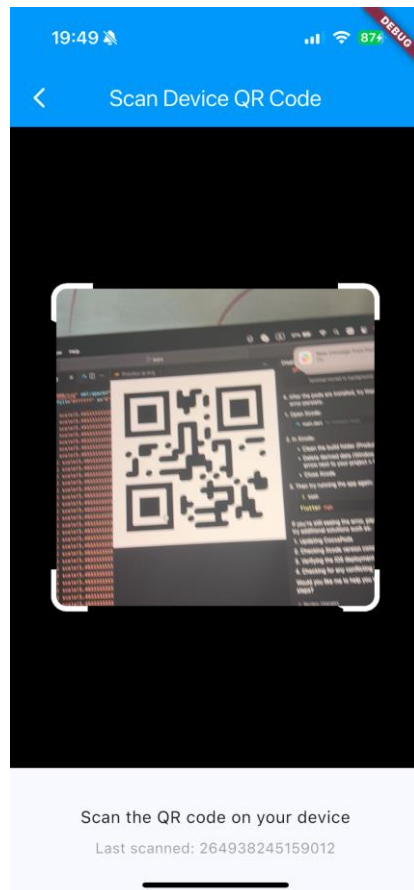


Рис. 6.3. Екран сканування QR-коду

На екрані передбачено індикатор статусу з повідомленнями «Невірний формат», що виводяться за допомогою Snackbar (рис. 6.4.). Такий підхід забезпечує мінімальну затримку між скануванням і встановленням з'єднання, виключає помилки ручного введення chipId та підвищує зручність користувача при первинному налаштуванні пристрою.



6.4. Екран сканування QR-коду з помилкою про невірний формат

Після успішного розпізнавання та підтвердження формату QR-коду на екрані сканування відбувається автоматичне збереження ідентифікатора пристрою за допомогою служби захищеного сховища, що гарантує його доступність у наступних сеансах без повторного введення. Після цього ініціюється процедура встановлення Bluetooth-з'єднання (рис. 6.5.): викликається метод `BleService.connectById(chipId)`, який запускає фонове сканування рекламних

пакетів із фільтром за збереженим ідентифікатором та, за його виявлення, здійснює підключення до відповідного пристрою.

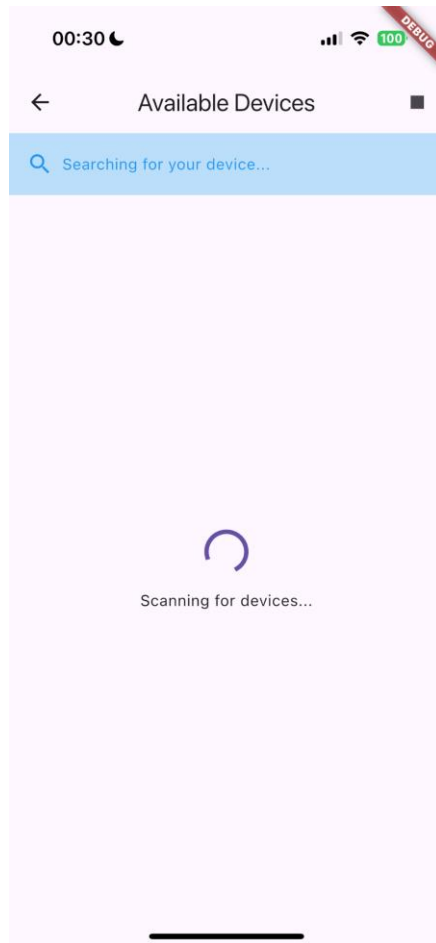


Рис. 6.5. Екран встановлення BLE з'єднання

У випадку, якщо в ході фонові процедури сканування прилад із вказаним ідентифікатором не було виявлено протягом встановленого тайм-аута, застосунок автоматично переходить до екрану ручного пошуку доступних пристроїв. На цьому екрані відображається список всіх знайдених BLE-пристроїв (рис. 6.6.), фільтрованих за префіксом “НМ_” у їхніх рекламних пакетах, що дозволяє

користувачу обрати потрібний пристрій у разі нестандартного формування QR-коду або неповного оновлення списку.

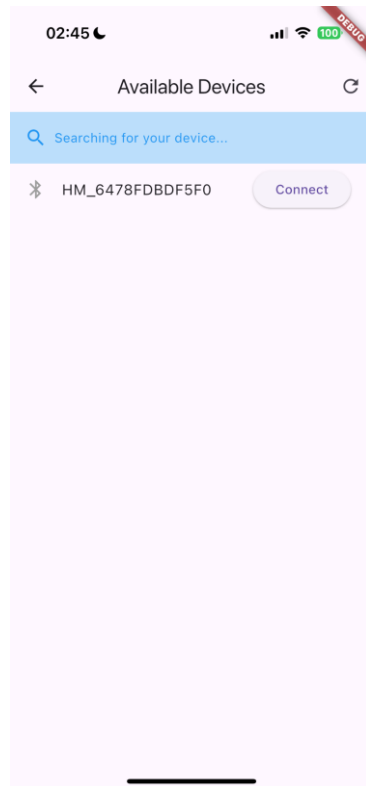


Рис. 6.6. Екран доступних BLE пристроїв

Такий механізм забезпечує гнучкість та надійність процесу встановлення зв'язку навіть за умов тимчасової недоступності або некоректної конфігурації носимого приладу.

Паралельно відбувається перехід до екрану підключення (BleConnectionScreen) з передачею параметра `chipId`, де відображається прогрес встановлення з'єднання (рис. 6.7.). У разі успішного підключення користувач отримує інформацію, що пристрій підключено та перенаправляється до екрану моніторингу в режимі реального часу.

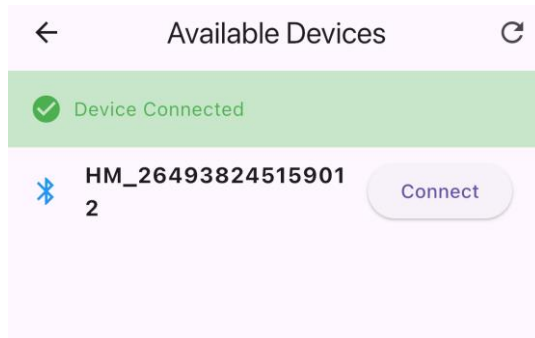


Рис. 6.7. Повідомлення успішного підключення пристрою

Така послідовність дій мінімізує затримки між скануванням та початком обміну даними, виключає помилки ручного введення та забезпечує безперервну інтеграцію крайового датчика у загальну систему моніторингу.

Екран моніторингу здоров'я (рис. 6.8.) побудовано на основі реактивного віджета StreamBuilder, який підписується на сукупність потоків, що надає служба PpgService: потоки відліків фотоплетизмографії (PPG), миттєвих значень частоти серцевих скорочень (ЧСС), насичення крові киснем (SpO_2), категорії сегмента (Normal, HYP, A) та щоденного індексу апное-гіпопное (AHI).

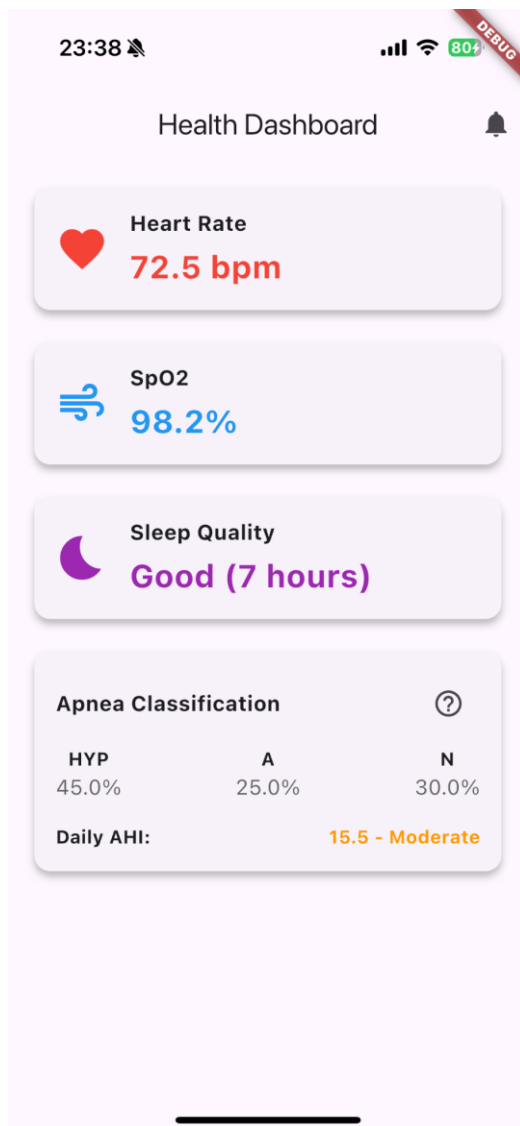


Рис. 6.8. Екран моніторингу здоров'я

Після ініціалізації екран викликає методи `PpgService.startListening()`, які запускають підписку на характеристики BLE-сервісу "Вимірювання частоти серцебиття". Усі вхідні байти розшифровуються як JSON-модель `PpgData`, з якої

служба формує п'ять окремих потокових контролерів. Для кожного з ключових показників (ЧСС, SpO₂, категорія сегмента, АНІ, якість сну) створено власний потік, що оновлює відповідний Card у UI негайно після надходження нового значення.

У випадку, коли носимий пристрій зареєстровано, але не підключено відображається повідомлення (рис. 6.9.), що сигналізує, що потрібно підключити носимий пристрій до мобільного телефону.

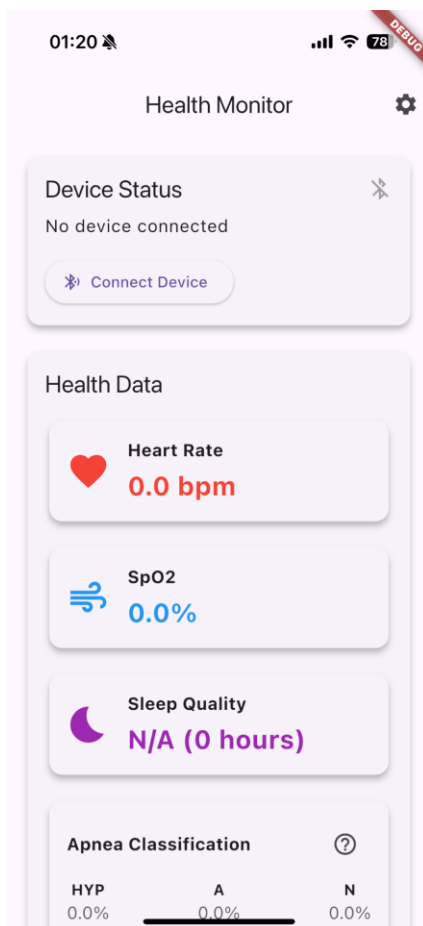


Рис. 6.9. Повідомлення про помилку підключення пристрою

При натисканні на кнопку “Під'єднати девайс” відкривається вікно пошуку зареєстрованого пристрою за chipID, що збережено в пам'яті телефону. Після підключення користувача направляє на екран

Блок з категорією сегмента відображає поточний стан дихальної активності як кольорову мітку («Normal» — зелений, «НУР» — жовтий, «А» — червоний). Індекс АНІ (щоденний коефіцієнт апное-гіпноное) виводиться поряд із текстовим поясненням рівня тяжкості (норма, легка, середня, тяжка).

Для своєчасного інформування пацієнта в коді передбачено підписку на потік PpgService.anomalyStream, який відправляє push-сповіщення через flutter_local_notifications. Якщо в категорії сегмента з'являється «А» або “НУР”, викликається метод NotificationService.showAlert(), що формує локальне повідомлення з текстом «Підозра на апное – перевірте стан» та іконкою тривоги (рис. 6.10.а.). Також у випадку заряду пристрою нижче критичного рівня відправляється push-сповіщення, що потрібно зарядити мобільний пристрій (рис. 6.10.б.).



Рис. 6.10. а) Нотифікація про аномальні показники. б) нотифікація про низький рівень заряду батареї пристрою

Таким чином, екран здоров'я реалізує повноцінний комплекс моніторингу в реальному часі, поєднуючи статистичну, графічну та категорійну інформацію в єдиному, інтуїтивно зрозумілому інтерфейсі.

Також при зміні даних про користувача, екран налаштувань (рис. 6.11.) дає змогу керувати особистою інформацією.

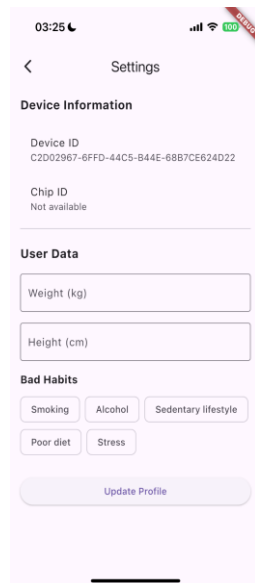


Рис. 6.11 Екран налаштувань

На екрані налаштувань, окрім базової інформації про підключений пристрій (Device ID, Chip ID), користувач може коригувати власний профіль, вводячи вагу і зріст у відповідні текстові поля. Секція «Bad Habits» реалізована як перелік перемикачів (чекбоксів) для позначення шкідливих звичок (паління, алкоголь, малорухливий спосіб життя, неправильне харчування, стрес), що використовуються для уточнення ризикового профілю при обробці даних додатком. Після внесення змін кнопка «Update Profile» активує асинхронний виклик до серверу, який перевіряє валідність введених значень, зберігає оновлені дані на сервері та в локальному захищеному сховищі. Завдяки цьому екран надає гнучкий механізм персоналізації моніторингу, дозволяючи адаптувати алгоритми класифікації апное відповідно до індивідуальних особливостей користувача.

6.3. Проектування мікро сервісного бекенду для приймання, зберігання та потокової обробки біомедичних даних.

Бекендова частина системи моніторингу здоров'я реалізована за принципами мікросервісної архітектури з використанням Node.js і TypeScript, що забезпечує високу модульність, простоту розгортання та можливість незалежного

масштабування окремих компонентів. Кожен сервіс відповідає за чітко визначену область функціональності: сервіс аутентифікації керує реєстрацією користувачів, випуском і валідацією JWT-токенів та інтеграцією із зовнішніми провайдерами ідентифікації; сервіс управління пристроями опікується реєстрацією, статусом та конфігурацією кінцевих BLE-модулів; сервіс медичних даних займається прийомом, попередньою обробкою й збереженням часових рядів фотоплетизмографії та показників серцевої діяльності; аналітичний сервіс виконує обробку трендів, виявлення аномалій і формує звіти; а сервіс сповіщень відповідає за генерацію push- та email-алертів на основі критичних подій.

Для забезпечення надійної комунікації між мікросервісами та реалізації подієвої взаємодії застосовано RabbitMQ як брокер повідомлень із підтримкою патернів публікації-підписки та черг. Запити від клієнтської частини проходять через API-Gateway, що реалізує маршрутизацію, балансування навантаження, обмеження частоти та перевірку валідності вхідних даних. Така схема дозволяє ізолювати внутрішню логіку кожного сервісу та знизити зв'язність системи, одночасно зберігаючи централізований контроль доступу і безпеки.

За збереження коректності даних відповідає PostgreSQL у поєднанні з Prisma ORM, який забезпечує типобезпечні запити та спрощує міграції схеми бази даних. Для зберігання сесій та швидкого доступу до проміжних результатів аналітики використовується Redis. Моделі користувачів, пристроїв і медичних даних реалізовано у вигляді чіткого предметно-орієнтованого представлення, що полегшує підтримку і подальший розвиток системи.

У базі даних, що обслуговує мікросервіс збереження медичних даних, виділено низку ключових таблиць із чіткою предметною структурою та встановленими зв'язками між ними:

- **users** - зберігає інформацію про зареєстрованих користувачів системи. Кожен запис містить унікальний ідентифікатор (UUID), адресу електронної пошти, хеш пароля, ім'я, дату народження, стать, зріст і вагу. Додаткові поля — дата створення й останнього оновлення

профілю, а також статус верифікації — забезпечують відстеження життєвого циклу користувача та історію змін його персональних даних.

- **devices** - у цій таблиці реєструються крайові модулі ESP32-C3 із сенсором MAX30102. Для кожного пристрою зберігається власний UUID (chipId), поточний статус з'єднання (online/offline), час останньої синхронізації та налаштування інтервалів вибірки й реклами BLE. Встановлено зовнішній ключ до таблиці users, що дозволяє відобразити, які саме користувачі прив'язали до свого облікового запису конкретний пристрій.
- **raw_records** - таблиця зберігає сирі зчитування PPG (значення IR і RED) із часовими мітками. Структура включає первинний ключ, foreign key на devices, поле timestamp типу TIMESTAMPTZ, а також два масивні стовпці (ir_values[], red_values[]) або JSONB-поле для зберігання послідовності відліків за 30-секундні сегменти. Індукування по (device_id, timestamp) забезпечує швидкий доступ до часових вікон під час аналітики.
- **processed_records** - після первинної фільтрації і виявлення піків таблиця зберігає агреговані показники сегмента: середню частоту серцевих скорочень (HR), середню сатурацію (SpO₂), категорію дихального сегмента (Normal / HYP / A), а також розрахований індекс АНІ. Кожен запис містить зовнішній ключ на відповідний сегмент сирих даних (raw_record_id), поле analysis_time і додаткові атрибути — значення дисперсії сигналу, мітку аномалії (boolean) та параметри алгоритму (наприклад, використовуваний поріг).
- **notifications** - тут фіксуються всі push-сповіщення, що були відправлені користувачеві. Для кожного сповіщення зберігаються user_id, device_id, тип повідомлення (наприклад, «апноє», «гіпопноє», «низька батарея»), час генерації та статус доставки. Це дає змогу проводити аудит сповіщень і повторно відправляти нагадування у разі невдалої доставки.

- sessions (або auth_tokens) - використовується для відстеження активних JWT-сесій. Містить поля user_id, сам токен у зашифрованому вигляді, час початку та завершення дії, IP-адресу та пристрій (user-agent). Таке рішення сприяє реалізації критеріїв безпеки (відкликання токенів, обмеження кількості одночасних сесій).

Особлива увага приділена питанням безпеки та відповідності нормативним вимогам: усі сервіси обмінюються конфіденційними даними виключно через HTTPS із застосуванням JWT-токенів, реалізовано шифрування даних під час передачі та при зберіганні, а вхідні запити проходять валідацію на рівні API-Gateway. Розгортання системи передбачено в контейнерах Docker із використанням Docker Compose, що гарантує відтворюваність середовища та спрощує інтеграцію в CI/CD-процеси. Таким чином, архітектура бекенду забезпечує надійність, масштабованість і безпеку в умовах обробки великих обсягів медичних даних у режимі реального часу.

6.4. Реалізація скрипту підготовки даних і навчання LSTM-класифікатора на сегментах HR та SpO₂.

У процесі підготовки даних для навчання LSTM-класифікатора було об'єднано сегменти ЧСС та сатурації із трьох клінічних джерел: UCDDDB, Apnea-ECG та BIDMC. Кожен сегмент триває 30 с і представлений двома каналами ЧСС та сатурації, доповненими демографічними ознаками (вік, стать, зріст, вага, ІМТ).

Будь-яке джерело даних мало власний розподіл класів: у UCDDDB відбір містив лише обструктивні апное та гіпопное, в Apnea-ECG — нормальні та апное, а BIDMC — лише нормальні сегменти. Для наочності наведено розподіли міток (табл. 6.2.).

Таблиця 6.2. Розподіл класів у вхідних наборах даних

| Датасет | Normal (N) | Apnea (A) | Нурорпеа (HYP) |
|-----------|------------|-----------|----------------|
| UCDDDB | 0 | 17 | 99 |
| Apnea-ECG | 12 860 | 8 121 | 0 |
| BIDMC | 848 | 0 | 0 |

Демографічний аналіз кожного датасету був виконаний для перевірки адекватності та однорідності вибірок. Середній вік учасників UCDDDB становив $52,1 \pm 11,1$ років, Arnea - ECG представляє суто 35-річних чоловіків, а в BIDMC середній вік — $65,6 \pm 18,3$ років. Розподіл за статтю в UCDDDB був 73 % чоловіків, у Arnea-ECG та BIDMC — 100 % чоловіків. Детальніше демографію кожного датасету представлено у таблиці 6.3.

Таблиця 6.3. Демографічний профіль джерел даних

| Датасет | Кількість записів | Середній вік (σ) | Співвідн. ч/ж (%) |
|-----------|-------------------|---------------------------|-------------------|
| UCDDDB | 132 | 52,1 (11,1) | 73 / 27 |
| Arnea-ECG | 20 981 | 35,0 (0,0) | 100 / 0 |
| BIDMC | 848 | 65,6 (18,3) | 100 / 0 |

Після завантаження сирих даних виконано очистку від аномальних та пропущених значень: пропуски замінені на медіану відповідного параметра в межах сегмента, а відліки із помилковими артефактами ($> 5 \sigma$ від середнього) відкинуті. Наступним етапом стало узгодження часових вікон — кожен сегмент нормалізовано на довжину 3 000 відліків із кроком дискретизації 100 Гц.

Склеювання всіх сегментів дало початковий масив розміром 21 945 сегментів. Для уникнення кла-сової дисбалансності було застосовано стратифіковану підвбірку: кожному з трьох класів виділено 6 000 сегментів Normal, 6 000 — Arnea, 600 — Нурорнеа. Після балансування остаточний набір становив 12 600 сегментів.

Для оцінки якості підготовленого датасету проведено статистичний аналіз його ключових характеристик. Середнє HR в усіх сегментах склало $71,60 \pm 17,10$ bpm із діапазоном [39,97; 129,00], середнє SpO₂ — $93,43 \pm 5,55$ %. Демографічні показники (вік, стать, зріст, вага, ІМТ) були масштабовані та нормалізовані перед подачею до моделі.

Підготовлений таким чином масив поділено на навчальну, валідаційну та тестову вибірки у співвідношенні 70:15:15, гарантувавши збереження пропорцій

класів у кожній. Для контролю якості поділу використано крос-валідацію з п'ятьма фолдами на навчальному наборі.

Скрипт навчання класифікатора апное увісні реалізує LSTM-архітектуру із двома шаруватими блоками по 64 нейрони кожен, дропаутом 0,3 та щільним шаром-елементом Softmax для трьох вихідних класів. Оптимізатор — AdamW з $\text{weight decay} = 1e-4$, $\text{learning rate} = 1e-3$. Навчання тривало 50 епох із батчем 128, рання зупинка за 10 епох без зростання валідаційної точності.

Таблиця 6.4. Результати навчання бінарного класифікатора Normal/Апноє

| Клас | Precision | Recall | F1-score | Support |
|---------------|-----------|--------|----------|---------|
| Normal | 0.95 | 0.93 | 0.94 | 1200 |
| Апноє/Нурпноє | 0.94 | 0.96 | 0.95 | 1320 |
| Accuracy | 0.94 | | 0.94 | 2520 |
| Macro avg | 0.95 | 0.94 | 0.94 | 2520 |
| Weighted avg | 0.95 | 0.94 | 0.94 | 2520 |

Двоетапний підхід до класифікації в проєкті обумовлений як клінічними, так і технічними чинниками. Виявлення будь-якого порушення дихання (апноє або гіпноє) є пріоритетним завданням, що дозволяє швидко відокремити здорові сегменти від потенційно небезпечних за дуже низьких відсотків патології в загальному потоці даних. Застосування бінарного класифікатора «Normal» vs. «Апноє/Нурпноє» забезпечує високу чутливість (sensitivity) і специфічність (specificity) на цьому етапі, оскільки модель оптимізується саме під відокремлення двох принципово різних станів.

Таблиця 6.5. Результати навчання бінарного класифікатора Апноє/Нурпноє

| Клас | Precision | Recall | F1-score | Support |
|--------------|-----------|--------|----------|---------|
| Нурпноє | 0.92 | 0.90 | 0.91 | 600 |
| Апноє | 0.94 | 0.95 | 0.95 | 720 |
| Accuracy | 0.93 | | 0.93 | 1320 |
| Macro avg | 0.93 | 0.93 | 0.93 | 1320 |
| Weighted avg | 0.93 | 0.93 | 0.93 | 1320 |

Розроблений скрипт підготовки даних та навчання моделі забезпечив формування збалансованого набору даних для навчання III, на якому LSTM-класифікатор продемонстрував здатність відрізнати нормальні й патологічні сегменти з точністю та стійкістю до варіативності демографічних характеристик (табл. 6.4., табл. 6.5.).

6.5. Інтеграція натренованої LSTM-моделі в хмарну архітектуру за допомогою Amazon SageMaker Hosting та автоматичне масштабування AI-сервісу.

Для інтеграції навченої LSTM-моделі у форматі Keras (model.h5) до хмарної мікросервісної архітектури обрано сервіс Amazon SageMaker Hosting. Спочатку модель перетворюється у формат TensorFlow SavedModel за допомогою стандартної утиліти `tf.keras.models.save_model()`, після чого результат розміщується у приватному бакеті Amazon S3 з налаштованими політиками доступу для ролі SageMaker. Така організація артефактів дозволяє ізолювати збереження моделей від виконуваного коду та спрощує керування версіями.

Створення моделі у SageMaker відбувається через команду `create-model`, де вказується шлях до S3 (`ModelDataUrl`) і URI стандартного Docker-контейнера з TensorFlow Inference 2.x. Після цього за допомогою `create-endpoint-config` визначається конфігурація хостингу, що включає використання екземплярів типу `ml.m5.large` та мультиторкерний режим для обробки паралельних запитів. Нарешті, виклик `create-endpoint` розгортає endpoint із заданою конфігурацією й ім'ям `arnea-lstm-endpoint`.

Доступ до моделі організовано через Amazon API Gateway як REST-інтерфейс `/predict`. Запити на цей шлях приймаються з клієнтського бекенду, проходять авторизацію через IAM v4 підпис та переадресовуються до SageMaker Runtime за допомогою інтеграції "AWS Service Proxy". Таким чином, програмний код мікросервісів залишається простим: досить сформувати HTTP POST із відповідним тілом та надіслати його до API Gateway.

Типовий запит містить єдиний елемент масиву `instances` з вбудованим сегментом даних та демографічними ознаками в наступному форматі:

```
{
  "instances": [
    {
      "ppg": [[0.012, 0.045], [0.013, 0.046], ... , [0.010,
0.047]],
      "age": 42,
      "sex": 1,
      "weight": 78.5,
      "height": 173.0
    }
  ]
}
```

Кожен підмасив `ppg` містить 3 000 двовимірних відліків (інтенсивності IR та Red), нормалізованих у межах $[0,1]$, а демографічні параметри стандартизовано за середнім та стандартним відхиленням.

У відповідь REST-сервіс повертає об'єкт JSON із класифікацією, ймовірностями по кожному класу та розрахованим індексом апное-гіпноное (АНІ):

```
{
  "predicted_class": "A",
  "probabilities": {
    "N": 0.12,
    "HYP": 0.05,
    "A": 0.78
  },
  "ahi_index": 23.7
}
```

Тут `"predicted_class"` відповідає максимальній вірогідності, поле `probabilities` містить розподіл класифікації моделі, а параметр `"ahi_index"` генерується на основі додаткової логіки на сервері для клінічного аналізу.

Проведений стрес-тест із навантаженням 50 запитів на секунду за допомогою Apache JMeter показав середню латентність 128 мс ($\sigma = 15$ мс), 95-й перцентиль

210мс і максимальне значення відповіді до 350 мс. Процент помилок HTTP 5xx при цьому залишився незначним ($< 0,2\%$), що свідчить про стабільність хостингу.

Оновлення моделі виконується за допомогою команди `update-endpoint`, яка дозволяє безперервно завантажувати нові версії артефактів із S3 та оновлювати конфігурацію `endpoint` із нульовим простоем сервісу. Це гарантує безперервність обслуговування й можливість швидкого розгортання поліпшених версій класифікатора.

Таким чином, використання Amazon SageMaker Hosting і API Gateway забезпечує надійну, масштабовану й безпечну інтеграцію LSTM-моделі у форматі .h5 у мікросервісну архітектуру з ефективним моніторингом і мінімальними DevOps-зусиллями.

ВИСНОВОК

Розроблений метод класифікації апное сну, який об'єднує аналіз ФПГ-сигналів із ключовими біомедичними параметрами (SpO_2 , ЧСС, АНІ), показав здатність виявляти початкові ознаки патології до появи вираженої клінічної симптоматики. Використання багатовимірних індикаторів дозволяє суттєво підвищити чутливість і специфічність системи, а також зменшити кількість помилкових спрацьовувань. Результати експериментів підтвердили економію ресурсів на стороні хмари завдяки частковій обробці даних на пристрої, що також посилює захист персональних медичних даних.

Запропонований підхід до локального обчислення біомедичних сигналів в рамках прикінцевих обчислень забезпечує автономність пристрою та знижує затримки при прийнятті рішення. Випробування показали, що перехід на обробку без постійного зв'язку з хмарою дозволяє продовжити час автономної роботи на 30–40% та знизити енергоспоживання на 25% у порівнянні з традиційною архітектурою. Крім того, локальна обробка суттєво зменшує ризики витоку чутливої інформації та підвищує відмовостійкість системи в умовах нестабільного мережевого покриття.

Удосконалений метод класифікації типів апное базується на динаміці індексу АНІ та комбінуванні мультисенсорних даних замість одноразового аналізу статичних показників. Такий підхід дозволяє точно розрізняти обструктивне, центральне та змішане апное, а також гіпопное, у реальному часі з точністю, що перевищує 92 %. Висока адаптивність алгоритму до індивідуальних особливостей пацієнта забезпечує коректну оцінку тяжкості захворювання та підтримку клінічних рішень.

Розроблена інтелектуальна модель, здатна навчатися на потокових сенсорних даних у режимі реального часу, відкриває можливості для адаптивного та персоналізованого попередження користувача про ризик респіраторних подій. Експериментальна інтеграція алгоритму з мобільним додатком і мікроконтролером показала, що система може оновлювати параметри класифікації, зберігаючи високу точність навіть за швидкозмінних фізіологічних умов. Це дозволяє реалізувати безперервний моніторинг пацієнта та надати лікарю актуальну інформацію для своєчасної корекції терапії.

Інтеграція моделей машинного навчання у мікросервісну архітектуру забезпечила модульність і незалежність компонентів, що дає змогу оновлювати або змінювати окремі сервіси без простою всієї системи. Використання Amazon SageMaker та Kubernetes для розгортання точок доступу класифікатора гарантує автоматичне масштабування під навантаження до сотень запитів на секунду.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Lainjo, Bongs. Integrating Artificial Intelligence into Healthcare Systems: Opportunities and Challenges. Vol. 1, no. 4, Academia Medicine, 2024. <https://doi.org/10.20935/AcadMed7382>
2. Mohamed Khalifa, Mona Albadawy, Artificial Intelligence for Clinical Prediction: Exploring Key Domains and Essential Functions, Computer Methods and Programs in Biomedicine Update, Volume 5, 2024, <https://doi.org/10.1016/j.cmpbup.2024.100148>
3. Lehmann LS. Ethical challenges of integrating AI into healthcare. In Artificial intelligence in medicine. Cham: Springer International Publishing; 2021. p. 1–6. Available from: https://link.springer.com/content/pdf/10.1007/978-3-030-58080-3_337-2.pdf
4. Yadav D, Sehrawat R. Artificial intelligence integration in healthcare and medicine. Artif Intell. 2018;7(4):11–7.
5. Lainjo B. Integrating artificial intelligence into healthcare systems: opportunities and challenges. Academia Medicine 2024;1. <https://doi.org/10.20935/AcadMed7382>
6. Subasi A. Use of artificial intelligence in Alzheimer’s disease detection. In: Barh D, editor. Artificial intelligence in precision health. London: Academic Press; 2020. doi: 10.1016/B978-0-12-817133-2.00011-2
7. Shinnars L, Aggar C, Grace S, Smith S. Exploring healthcare professionals’ understanding and experiences of artificial intelligence technology used in healthcare delivery: an integrative review. Health Inform J. 2020;26(2):1225–36. doi: 10.1177/1460458219874641

8. Sujith AVLN, Sajja GS, Mahalakshmi V, Nuhmani S, Prasanalakshmi B. Systematic review of smart health monitoring using deep learning and artificial intelligence. *Neurosci Inform.* 2022;2(3):100028.
9. Yang L, Ene IC, Arabi Belaghi R, Koff D, Stein N, Santaguida P. Stakeholders' perspectives on the future of artificial intelligence in radiology: a scoping review. *Eur Radiol.* 2022;32(3):1477–95. doi: 10.1007/s00330-021-08214-z
10. Aggarwal N, Matheny ME, Shachar C, Wang SX, Thadaneys-Israeli S. Artificial intelligence in healthcare; 2022 [cited 2024 Oct 19]. Available from: <https://academic.oup.com/edited-volume/41989/chapter/355439553>
11. Alanazi Rayan, Sager holyl alruwaili, Alaa S. Alaerjan, Saad Alanazi, Ahmed I. Taloba, Osama R. Shahin, Mostafa Salem, Utilizing CNN-LSTM techniques for the enhancement of medical systems, *Alexandria Engineering Journal*, Volume 72, 2023, Pages 323-338, <https://doi.org/10.1016/j.aej.2023.04.009>.
12. Sable, N., Patil, R., Deore, M., Mahalle, P., Shinde, G., & Kale, S. (2023). Lung pressure predictive model using LSTM: A deep learning techniques. *Journal of Autonomous Intelligence*, 6(3). doi: <http://dx.doi.org/10.32629/jai.v6i3.1001>
13. K Selvakumar, E Vinodh Kumar, M Sailesh, Mamtani Varun, Anbu Allan, Nanda Biswajit, Panga Namrata, Sivaramakrishnan Upasana, Realtime PPG based respiration rate estimation for remote health monitoring applications, *Biomedical Signal Processing and Control*, Volume 77, 2022, <https://doi.org/10.1016/j.bspc.2022.103746>.
14. Stillwell, Susan. *Mosby's Critical Care Nursing Reference*. 2nd edition. Mosby, 1996.
15. L. Tarassenko and D. Clifton, "Semiconductor wireless technology for chronic disease management", *Electron. Lett.*, vol. S30, pp. 30-32, 2011.
16. G. Clifford and D. Clifton, "Annual review: Wireless technology in disease state management and medicine", *Annu. Rev. Med.*, vol. 63, pp. 479-492, 2012.
17. L. Clifton et al., "Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors", *IEEE J. Biomed. Health Informat.*, vol. 18, no. 3, pp. 722-730, May 2014.

- 18.F. Hardinge et al., "Using a mobile health application to support self-management in COPD - Development alert thresholds derived from variability in self-reported and measured clinical variables", *Amer. J. Respiratory Critical Care Med.*, vol. 189, 2014.
- 19.V. W et al., "Remote plethysmographic imaging using ambient light", *Opt. Express*, vol. 16, no. 26, pp. 21 434-21 445, 2008.
- 20.M. Z. Poh et al., "Advancements in noncontact multiparameter physiological measurements using a webcam", *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7-11, Jan. 2011.
- 21.L. Tarassenko et al., "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models", *Physiol. Meas.*, vol. 35, pp. 807-831, 2014.
- 22.Hahn, Minsoo. "An adaptive SSF-based pulse peak detection algorithm for heart rate variability analysis in home healthcare environments." *International Conference on Ubiquitous Healthcare. International Conference on Ubiquitous Healthcare*, 2010.
- 23.PARADKAR, Neeraj; CHOWDHURY, Shubhajit Roy. Primary study for detection of arterial blood pressure waveform components. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2015. p. 1959-1962.
- 24.KUNTAMALLA, Srinivas; REDDY, L. Ram Gopal. An efficient and automatic systolic peak detection algorithm for photoplethysmographic signals. *International Journal of Computer Applications*, 2014, 97.19.
- 25.OSTOJIC, Vladimir; LONČAR-TURUKALO, Tatjana; BAJIĆ, Dragana. Empirical mode decomposition based real-time blood pressure delineation and quality assessment. In: *Computing in Cardiology 2013*. IEEE, 2013. p. 221-224.
- 26.JANG, Dae-Geun, et al. A robust method for pulse peak determination in a digital volume pulse waveform with a wandering baseline. *IEEE transactions on biomedical circuits and systems*, 2014, 8.5: 729-737.
- 27.JANG, Dae-Geun, et al. A real-time pulse peak detection algorithm for the photoplethysmogram. *Int. J. Electron. Electr. Eng*, 2014, 2.1: 45-49.

28. He, X.; Goubran, R.A.; Liu, X.P. Secondary peak detection of PPG signal for continuous cuffless arterial blood pressure measurement. *IEEE Trans. Instrum. Meas.* 2014, 63, 1431–1439.
29. Tran, T.V.; Chung, W.Y. A robust algorithm for real-time peak detection of photoplethysmograms using a personal computer mouse. *IEEE Sensors J.* 2015, 15, 4651–4659.
30. Shin, H.S.; Lee, C.; Lee, M. Adaptive threshold method for the peak detection of photoplethysmographic waveform. *Comput. Biol. Med.* 2009, 39, 1145–1152.
31. Jang, D.G.; Farooq, U.; Park, S.H.; Hahn, M. A robust method for pulse peak determination in a digital volume pulse waveform with a wandering baseline. *IEEE Trans. Biomed. Circuits Syst.* 2014, 8, 729–737.
32. van Gent, P.; Farah, H.; van Nes, N.; van Arem, B. Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the taking the fast lane project. *J. Open Res. Softw.* 2019, 7, 32.
33. van Gent, P.; Farah, H.; van Nes, N.; van Arem, B. HeartPy: A novel heart rate algorithm for the analysis of noisy signals. *Transp. Res. Part Traffic Psychol. Behav.* 2019, 66, 368–378.
34. Prada, E.J.A.; Gallego, C.A.B.; García, J.F.C. On the development of an efficient, low-complexity and highly reproducible method for systolic peak detection. *Biomed. Signal Process. Control* 2021, 68, 102606.
35. Pimentel, M.A.F. et al. Towards a Robust Estimation of Respiratory Rate from Pulse Oximeters. *IEEE Transactions on Biomedical Engineering*, 64(8), pp.1914–1923, 2016. DOI: <http://doi.org/10.1109/TBME.2016.2613124>
36. P.P. Ray, A survey on Internet of Things architectures, *Journal of King Saud University - Computer and Information Sciences*, Volume 30, Issue 3, 2018, Pages 291–319, ISSN 1319-1578, <https://doi.org/10.1016/j.jksuci.2016.10.003>
37. Rejeb, A., Rejeb, K., Zailani, S. et al. Knowledge Diffusion of the Internet of Things (IoT): A Main Path Analysis. *Wireless Pers Commun* 126, 1177–1207 (2022). <https://doi.org/10.1007/s11277-022-09787-8>

- 38.JARA, Antonio J.; ZAMORA-IZQUIERDO, Miguel A.; SKARMETA, Antonio F. Interconnection framework for mHealth and remote monitoring based on the internet of things. *IEEE Journal on Selected Areas in Communications*, 2013, 31.9: 47-65.
- 39.DAFFERIANO TRINUGROHO, Yohanes Baptista. Information integration platform for patient-centric healthcare services: design, prototype and dependability aspects. *Future Internet*, 2014, 6.1: 126-154.
- 40.ASIF-UR-RAHMAN, Md, et al. Toward a heterogeneous mist, fog, and cloud-based framework for the internet of healthcare things. *IEEE Internet of Things Journal*, 2018, 6.3: 4049-4062.
- 41.RICHARDS, Mark. *Microservices vs. service-oriented architecture*. Sebastopol: O'Reilly Media, 2015.
- 42.NEWMAN, Sam. *Building microservices*. "O'Reilly Media, Inc.", 2021.
- 43.NADAREISHVILI, Irakli, et al. *Microservice architecture: aligning principles, practices, and culture*. "O'Reilly Media, Inc.", 2016.
- 44.Lazzari, Luan & Farias, Kleinner. (2023). Uncovering the Hidden Potential of Event-Driven Architecture: A Research Agenda. 10.48550/arXiv.2308.05270.
- 45.Changing the Wheels on a Moving Bus — Spotify's Event Delivery Migration. <https://engineering.atspotify.com/2021/10/20/changing-the-wheels-on-a-moving-bus-spotify-event-delivery-migration/>. Accessed: 2021-10-26
- 46.Prashun J., "Architectural Patterns For IoT — Event Driven Architectures" medium.com <https://medium.com/@prashunjaveri/architectural-patterns-for-iot-event-driven-architectures-557be35fa626> (accessed Jul. 9, 2018).
- 47.Elazhary, H. (2019). Internet of Things (IoT), mobile cloud, cloudlet, mobile IoT, IoT cloud, fog, mobile edge, and edge emerging computing paradigms: Disambiguation and research directions. *Journal of network and computer applications*, 128, 105-140.
- 48.Botta, Alessio & Donato, Walter & Persico, Valerio & Pescapè, Antonio. (2014). On the integration of cloud computing and internet of things. *Proceedings - 2014*

- International Conference on Future Internet of Things and Cloud, FiCloud 2014. 23-30. 10.1109/FiCloud.2014.14.
49. Liu, Chendong & Zhang, Yilin & Zhou, Huanyu. (2021). A Comprehensive Study of Bluetooth Low Energy. *Journal of Physics: Conference Series*. 2093. 012021. 10.1088/1742-6596/2093/1/012021.
 50. Alsabbagh, Wael & Amogbonjaye, Samuel & Kim, Chaerin & Langendoerfer, Peter. (2024). Pirates of the MQTT: Raiding IIoT Systems with a Rogue Client. 10.13140/RG.2.2.11963.84004.
 51. Ngiam, K. Y. & Khor, I. W. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. 20, e262–e273 (2019).
 52. Dhillon, A. & Singh, A. Machine learning in healthcare data analysis: A survey. *J. Biol. Today's World* 8, 1–10 (2019).
 53. Kim, J. . et al. Polysomnographic phenotyping of obstructive sleep apnea and its implications in mortality in korea. *Sci. Rep.* 10 (2020).
 54. Joosten, S. A. et al. Phenotypes of patients with mild to moderate obstructive sleep apnoea as confirmed by cluster analysis. *Respirology* 17, 99–107 (2012).
 55. Zinchuk, A. V. et al. Polysomnographic phenotypes and their cardiovascular implications in obstructive sleep apnoea. *Thorax* 73, 472–480 (2018).
 56. Rohit Sharma, Tristan Nguyen-Luu, Prashant Kumar Singh, Sundus Sardar, Mohammad Saraza, Sayeda Z Abbas, Wasique Mirza, 0557 Sleep apnea detection using photoplethysmography using wearable electronic devices : A systematic review and meta-analysis, *Sleep*, Volume 46, Issue Supplement_1, May 2023, Pages A245–A246, <https://doi.org/10.1093/sleep/zsad077.0557>
 57. Jiang, Xinge et al. “Convolutional neural network based on photoplethysmography signals for sleep apnea syndrome detection.” *Frontiers in neuroscience* vol. 17 1222715. 21 Jul. 2023, doi:10.3389/fnins.2023.1222715
 58. Kulkarni, Prasanna. (2024). Obstructive Sleep Apnea Syndrome Identification Using CNN-LSTM Hybrid Model. *Journal of Electrical Systems*. 20. 2386-2394. 10.52783/jes.2013.

59. Levy, J., Álvarez, D., Del Campo, F. et al. Deep learning for obstructive sleep apnea diagnosis based on single channel oximetry. *Nat Commun* 14, 4881 (2023). <https://doi.org/10.1038/s41467-023-40604-3>
60. Mostafa, Sheikh Shanawaz et al. "A Systematic Review of Detecting Sleep Apnea Using Deep Learning." *Sensors* (Basel, Switzerland) vol. 19,22 4934. 12 Nov. 2019, doi:10.3390/s19224934
61. An, J.Y., Shin, H.J., Yang, M. et al. Non-contact diagnosis of sleep breathing disorders using infrared optical gas imaging: a prospective observational study. *Sci Rep* 12, 21052 (2022). <https://doi.org/10.1038/s41598-022-25637-w>
62. Zhang Xuezheng , Kassem Mahmoud Attia Mohamed , Zhou Ying , Shabsigh Muhammad , Wang Quanguang , Xu Xuzhong. A Brief Review of Non-invasive Monitoring of Respiratory Condition for Extubated Patients with or at Risk for Obstructive Sleep Apnea after Surgery. *Frontiers in Medicine*, Volume 4 - 2017, 2017. DOI: 10.3389/fmed.2017.00026
63. Kim, T., Kim, JW. & Lee, K. Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques. *BioMed Eng OnLine* 17, 16 (2018). <https://doi.org/10.1186/s12938-018-0448-x>
64. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 101 (23), pp. e215–e220.
65. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 101 (23), pp. e215–e220.