

Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна

Навчально-науковий інститут комп'ютерних наук та штучного інтелекту

Кафедра математичного моделювання та аналізу даних

До захисту допущено

Кафедрою ММтАД

протокол №

від

завідувач кафедри

(підпис)

Володимир Струков

(ім'я, прізвище)

« » 2026 р.

Кваліфікаційна робота
здобувача першого (бакалаврського) рівня вищої освіти

Оцінка біометричних показників людини з використанням глибокого навчання
(назва роботи)

Спеціальність (спеціалізація) 122 Комп'ютерні науки

Освітня програма Комп'ютерні науки

Виконавець


(підпис)

Ігор Загоруйко

(ім'я, прізвище)

Науковий керівник


(підпис)

Володимир Донець

(ім'я, прізвище)

Харків – 2026

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний університет імені В. Н. Каразіна

Навчально-науковий інститут комп'ютерних наук та штучного інтелекту
Кафедра Математичного моделювання та аналізу даних
Рівень вищої освіти (освітньо-кваліфікаційний рівень) Бакалавр
Галузь знань: F – Інформаційні технології
Спеціальність: F3 – Комп'ютерні науки
Освітня програма: Комп'ютерні науки

ЗАТВЕРДЖУЮ
Завідувач кафедри ММ та АД
Струков Володимир Михайлович
« 13 » жовтня 2025 року

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ (ПРОЕКТ)

Загоруйко Ігор Васильович
(прізвище, ім'я, по батькові студента)

1. Тема роботи Оцінка біометричних показників людини з використанням
глибокого навчання

керівник роботи Донець Володимир Віталійович,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від “ 29 ” січня 2026 року № **4101-5/225**

2. Строк подання студентом роботи 01 червня 2026 року

3. Перелік питань, які потрібно розробити:

- 1) проаналізувати предметну область оцінювання біометричних показників людини за фізіологічними сигналами;
- 2) розглянути особливості фотоплетизмографічного сигналу та його використання для оцінювання частоти серцевих скорочень;
- 3) проаналізувати класичні методи оцінювання ЧСС за PPG-сигналом і підходи на основі машинного та глибокого навчання;
- 4) вибрати джерело даних і релевантні канали фізіологічних сигналів для проведення експериментів;
- 5) розробити процедуру формування набору даних із 30-секундних PPG-вікон і відповідних референсних значень ЧСС;

- 6) забезпечити коректне розбиття даних на навчальну, валідаційну й тестову вибірки без витоку між клінічними кейсами;
- 7) реалізувати попередню обробку PPG-сигналу: заповнення пропусків, фільтрацію, зменшення частоти дискретизації та нормалізацію;
- 8) реалізувати базовий статистичний метод оцінювання ЧСС на основі пошуку піків PPG-сигналу;
- 9) побудувати й навчити моделі XGBoost, Conv+LSTM та Conv+Transformer для оцінювання ЧСС;
- 10) провести порівняльне експериментальне оцінювання моделей, проаналізувати похибки, обмеження та сформулювати висновки.

4. План роботи

№ з/п	Назви етапів роботи	Термін виконання етапів роботи
1	Уточнення теми, мети, об'єкта, предмета та загальної постановки задачі дослідження.	13.10.2025–26.10.2025
2	Огляд предметної області, біометричних показників, PPG-сигналу та методів оцінювання ЧСС.	27.10.2025–09.11.2025
3	Аналіз класичних методів, підходів машинного навчання та глибокого навчання для фізіологічних часових рядів.	10.11.2025–23.11.2025
4	Вибір відкритого джерела даних VitalDB і релевантних каналів SNUADC/PLETH та Solar8000/HR.	24.11.2025–07.12.2025
5	Розроблення алгоритму формування 30-секундних PPG-вікон і підготовки референсних значень ЧСС.	08.12.2025–28.12.2025
6	Реалізація розбиття даних на навчальну, валідаційну й тестову вибірки за клінічними кейсами.	29.12.2025–11.01.2026
7	Розроблення попередньої обробки PPG-сигналу та показників якості PPG-вікон.	12.01.2026–01.02.2026
8	Реалізація статистичного базового методу оцінювання ЧСС на основі піків PPG-сигналу.	02.02.2026–15.02.2026
9	Побудова моделі XGBoost на основі ручних ознак PPG-вікна.	16.02.2026–01.03.2026
10	Проектування, навчання та валідація компактних нейронних моделей Conv+LSTM і Conv+Transformer.	02.03.2026–29.03.2026
11	Порівняльне тестування моделей за регресійними метриками та аналіз похибок за клінічними кейсами, діапазонами ЧСС і якістю сигналу.	30.03.2026–19.04.2026
12	Оформлення пояснювальної записки, додатків, висновків і підготовка матеріалів до подання роботи.	20.04.2026–01.06.2026

5. Дата видачі завдання 13.10.2025

Студент Загоруйко Ігор Васильович
ініціали, прізвище
підпис



Керівник роботи Донець Володимир Віталійович
ініціали, прізвище
підпис



АНОТАЦІЯ

Кваліфікаційна робота містить 94 сторінки, 3 розділи, 13 таблиць, 10 рисунків, 1 додаток, 20 джерел посилання.

Кваліфікаційна робота присвячена дослідженню, програмній реалізації та експериментальному оцінюванню методів оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом із використанням методів машинного та глибокого навчання. У роботі розглянуто задачу автоматизованого оцінювання частоти серцевих скорочень за 30-секундними фрагментами сигналу фотоплетизмографії, отриманими з клінічних записів VitalDB.

Метою роботи є підвищення точності та стійкості автоматизованого оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом порівняно з класичним статистичним методом шляхом розроблення й експериментального оцінювання програмного підходу на основі методів машинного та глибокого навчання. Об'єктом дослідження є фізіологічні часові ряди, що характеризують біометричні показники людини. Предметом дослідження є методи попередньої обробки фотоплетизмографічного сигналу та моделі машинного і глибокого навчання для оцінювання частоти серцевих скорочень.

У роботі використано відкриті клінічні дані VitalDB. Вхідним сигналом є канал SNUADC/PLETH, а референсним значенням частоти серцевих скорочень – канал Solar8000/HR. Було сформовано набір даних із 6157 клінічних кейсів, поділених на навчальну, валідаційну й тестову вибірки за ідентифікатором клінічного запису. Такий спосіб поділу запобігає витоку даних між вибірками, оскільки вікна одного клінічного запису не потрапляють одночасно до навчання й тестування. Загалом сформовано 2129005 вікон тривалістю 30 секунд.

У межах роботи реалізовано попередню обробку сигналу, що включає заповнення пропусків, смугову фільтрацію, зменшення частоти дискретизації та стійку нормалізацію. Для оцінювання якості сигналу розраховано частку валідних відліків, міжквартильний розмах амплітуди, стандартне відхилення

підготовленого сигналу та частку потенційно обмежених значень. Для прогнозування частоти серцевих скорочень порівняно чотири підходи: статистичний метод на основі пошуку піків фотоплетизмографічного сигналу, XGBoost на основі ручних ознак, Conv+LSTM та Conv+Transformer.

Найкращі результати показала модель Conv+LSTM. На спільному тестовому наборі з 318268 вікон вона досягла середньої абсолютної похибки 1,803 уд/хв, кореня середньоквадратичної похибки 6,075 уд/хв і коефіцієнта детермінації 0,842. Частка прогнозів із похибкою не більше 5 уд/хв становила 93,64%, а з похибкою не більше 10 уд/хв – 96,59%. Аналіз на рівні клінічних кейсів підтвердив стабільність моделі: медіана середньої абсолютної похибки за кейсами становила 1,142 уд/хв, а 95-й перцентиль – 5,885 уд/хв.

Практичне значення роботи полягає у створенні відтворюваного програмного підходу до обробки фотоплетизмографічних сигналів і оцінювання частоти серцевих скорочень. Отримані результати демонструють перспективність компактних нейронних моделей для аналізу біометричних показників людини. Основним обмеженням роботи є погіршення якості на рідкісних високих та екстремальних значеннях частоти серцевих скорочень, а також наявність артефактів у референсному каналі.

КЛЮЧОВІ СЛОВА: БІОМЕТРИЧНІ ПОКАЗНИКИ, ЧАСТОТА СЕРЦЕВИХ СКОРОЧЕНЬ, ФОТОПЛЕТИЗМОГРАФІЯ, PPG, ГЛИБОКЕ НАВЧАННЯ, МАШИННЕ НАВЧАННЯ, ЧАСОВІ РЯДИ, VITALDB, XGBOOST, LSTM, TRANSFORMER.

ABSTRACT

The qualification thesis contains 94 pages, 3 chapters, 13 tables, 10 figures, 1 appendix, and 20 references.

The qualification thesis is devoted to the research, software implementation, and experimental evaluation of methods for estimating a human biometric indicator, namely heart rate, from a photoplethysmographic signal using machine learning and deep learning methods. The work considers the task of automated heart rate estimation from 30-second photoplethysmographic signal segments obtained from VitalDB clinical records.

The purpose of the work is to improve the accuracy and robustness of automated heart rate estimation from a photoplethysmographic signal compared with a classical statistical method by developing and experimentally evaluating a software approach based on machine learning and deep learning methods. The object of the research is physiological time series describing human biometric indicators. The subject of the research is photoplethysmographic signal preprocessing methods and machine learning and deep learning models for heart rate estimation.

The study uses open clinical data from VitalDB. The input signal is the SNUADC/PLETH channel, while the reference heart rate is taken from the Solar8000/HR channel. A dataset of 6157 clinical cases was formed and split into training, validation, and test subsets by clinical case identifier. This splitting strategy prevents data leakage between subsets, since windows from the same clinical record do not appear simultaneously in training and testing. In total, 2129005 windows with a duration of 30 seconds were generated.

The implemented preprocessing procedure includes missing value filling, band-pass filtering, downsampling, and robust normalization. Signal quality indicators were also calculated, including the proportion of valid samples, amplitude interquartile range, standard deviation of the prepared signal, and the proportion of potentially clipped values. Four approaches were compared for heart rate estimation: a statistical

method based on photoplethysmographic peak detection, XGBoost based on handcrafted features, Conv+LSTM, and Conv+Transformer.

The best results were achieved by the Conv+LSTM model. On the common test subset of 318268 windows, it achieved a mean absolute error of 1.803 bpm, a root mean squared error of 6.075 bpm, and a coefficient of determination of 0.842. The proportion of predictions with an absolute error not exceeding 5 bpm was 93.64%, and the proportion not exceeding 10 bpm was 96.59%. Case-level analysis confirmed the stability of the model: the median case-level mean absolute error was 1.142 bpm, while the 95th percentile was 5.885 bpm.

The practical value of the work lies in the development of a reproducible software approach for photoplethysmographic signal processing and heart rate estimation. The obtained results demonstrate the potential of compact neural network models for the analysis of human biometric indicators. The main limitation of the work is reduced performance on rare high and extreme heart rate values, as well as the presence of artifacts in the reference channel.

KEYWORDS: BIOMETRIC INDICATORS, HEART RATE, PHOTOPLETHYSMOGRAPHY, PPG, DEEP LEARNING, MACHINE LEARNING, TIME SERIES, VITALDB, XGBOOST, LSTM, TRANSFORMER.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ І ТЕРМІНІВ	11
ВСТУП.....	13
РОЗДІЛ 1. ОГЛЯД І АНАЛІЗ МЕТОДІВ ОЦІНЮВАННЯ БІОМЕТРИЧНИХ ПОКАЗНИКІВ ЛЮДИНИ ЗА ФІЗІОЛОГІЧНИМИ СИГНАЛАМИ.....	18
1.1. Біометричні показники людини як об'єкт комп'ютерного аналізу	18
1.2. Частота серцевих скорочень як біометричний і фізіологічний показник.....	20
1.3. Фотоплетизмографічний сигнал: природа, структура та особливості	21
1.4. Класичні методи оцінювання частоти серцевих скорочень за PPG- сигналом	22
1.5. Методи машинного навчання в задачах аналізу фізіологічних сигналів.....	24
1.6. Глибоке навчання для обробки одновимірних часових рядів.....	25
1.7. Набори даних для дослідження фізіологічних сигналів.....	27
1.8. Проблеми оцінювання якості моделей у задачах аналізу медичних часових рядів	28
1.9. Мотивований висновок щодо необхідності виконання роботи	29
1.10. Висновки до розділу 1	30
РОЗДІЛ 2. ПОСТАНОВКА ЗАДАЧІ, МЕТОДИКА ТА ПРОГРАМНА РЕАЛІЗАЦІЯ ОЦІНЮВАННЯ ЧАСТОТИ СЕРЦЕВИХ СКОРОЧЕНЬ ЗА PPG- СИГНАЛОМ.....	32
2.1. Постановка задачі дослідження	32
2.2. Об'єкт, предмет, мета та завдання роботи	33

2.3. Обмеження задачі та припущення дослідження	34
2.4. Загальна декомпозиція програмного рішення	35
2.5. Вибір джерела даних і каналів фізіологічних сигналів	37
2.6. Формування набору даних із 30-секундних вікон	38
2.7. Розбиття даних на навчальну, валідаційну та тестову вибірки.....	42
2.8. Попередня обробка фотоплетизмографічного сигналу	43
2.9. Розрахунок показників якості PPG-вікон.....	44
2.10. Статистичний метод оцінювання ЧСС на основі піків PPG	45
2.11. Модель XGBoost	46
2.12. Архітектура нейронної моделі Conv+LSTM	49
2.13. Архітектура нейронної моделі Conv+Transformer	51
2.14. Організація навчання моделей	53
2.15. Метрики оцінювання якості прогнозування	54
2.16. Оцінювання невизначеності та аналіз на рівні клінічних кейсів ..	55
2.17. Висновки до розділу 2	56

РОЗДІЛ 3. ЕКСПЕРИМЕНТАЛЬНІ РЕЗУЛЬТАТИ ТА АНАЛІЗ ЯКОСТІ РОЗРОБЛЕНИХ МОДЕЛЕЙ

58

3.1. Характеристика сформованого набору даних.....	58
3.2. Розподіл референсної частоти серцевих скорочень	60
3.3. Основне порівняння моделей на спільній тестовій вибірці	61
3.4. Аналіз розподілу абсолютної похибки	64
3.5. Аналіз результатів на рівні клінічних кейсів	65
3.6. Аналіз узгодженості прогнозів методом Бленда-Альтмана	67
3.7. Калібраційний аналіз прогнозованої та референсної ЧСС.....	68

3.8. Аналіз похибки за діапазонами ЧСС	70
3.9. Аналіз впливу якості PPG-сигналу на похибку	71
3.10. Аналіз кореляції помилок між моделями	73
3.11. Оцінювання швидкодії та компактності моделей.....	75
3.12. Аналіз результатів після фільтрації фізіологічно правдоподібного діапазону ЧСС	76
3.13. Узагальнення отриманих результатів і практичні обмеження	76
3.14. Висновки до розділу 3	78
ВИСНОВКИ	81
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	85
ДОДАТОК А. ФРАГМЕНТИ ПРОГРАМНОЇ РЕАЛІЗАЦІЇ.....	88
Лістинг А.1 – Загальна логіка формування 30-секундних вікон	88
Лістинг А.2 – Попередня обробка PPG-сигналу.....	90
Лістинг А.3 – Обчислення показників якості PPG-вікна	91
Лістинг А.4 – Архітектура Conv+LSTM	92
Лістинг А.5 – Обчислення основних метрик	94

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ І ТЕРМІНІВ

Аналіз Бленда-Альтмана – метод оцінювання узгодженості двох способів вимірювання або прогнозування.

Бутстреп – статистичний метод оцінювання невизначеності шляхом багаторазового повторного вибирання елементів із вибірки.

Бутстреп за клінічними кейсами – бутстреп, у якому одиницею повторного вибирання є клінічний кейс, а не окреме вікно.

Валідаційна вибірка – частина даних, використана для вибору параметрів і контролю якості моделей під час навчання.

Вікно сигналу – фрагмент PPG-сигналу фіксованої тривалості; у цій роботі – 30 секунд.

Витік даних – ситуація, коли інформація з тестової вибірки прямо або опосередковано потрапляє до процесу навчання моделі, що призводить до завищення оцінки якості.

Зсув, bias – середній зсув прогнозу відносно референсного значення.

Клінічний кейс – окремий клінічний запис у VitalDB, що містить часові ряди фізіологічних сигналів і моніторних показників.

Міжквартильний розмах амплітуди PPG-сигналу – різниця між 75-м і 25-м процентилями амплітуди сирого PPG-сигналу.

Навчальна вибірка – частина даних, використана для навчання моделей.

Стандартне відхилення підготовленого PPG-сигналу – показник варіативності сигналу після заповнення пропусків, фільтрації, зменшення частоти дискретизації та нормалізації.

Тестова вибірка – частина даних, використана лише для фінального оцінювання якості моделей.

Частка валідних відліків PPG-сигналу – відношення кількості скінченних і придатних для обробки відліків до загальної кількості відліків у сирому PPG-вікні.

Частка потенційно обмежених за амплітудою значень – частка відліків PPG-сигналу, що знаходяться поблизу крайніх рівнів амплітуди та можуть свідчити про насичення або обрізання сигналу.

ЧСС – частота серцевих скорочень, кількість серцевих скорочень за одну хвилину.

AE – absolute error, абсолютна похибка.

Conv+LSTM – нейронна модель, що поєднує одновимірні згорткові шари та рекурентний шар LSTM.

Conv+Transformer – нейронна модель, що поєднує згорткове перетворення сигналу та кодувальник Transformer.

HR – heart rate, частота серцевих скорочень.

LSTM – long short-term memory, різновид рекурентної нейронної мережі.

MAE – mean absolute error, середня абсолютна похибка.

P90 AE – 90-й перцентиль абсолютної похибки.

P95 AE – 95-й перцентиль абсолютної похибки.

PPG – photoplethysmography, фотоплетизмографія; метод оптичної реєстрації змін кровонаповнення тканин.

R2 – коефіцієнт детермінації.

RMSE – root mean squared error, корінь середньоквадратичної похибки.

Solar8000/HR – канал моніторної частоти серцевих скорочень у даних VitalDB.

SNUADC/PLETH – канал фотоплетизмографічного сигналу в даних VitalDB.

Transformer – архітектура нейронної мережі на основі механізму уваги.

VitalDB – відкритий набір клінічних фізіологічних даних, використаний у роботі як джерело записів PPG та референсної ЧСС.

XGBoost – алгоритм градієнтного бустингу дерев рішень, використаний як модель машинного навчання на основі ручних ознак.

ВСТУП

Біометричні показники людини відіграють важливу роль у сучасних інформаційних системах, медичному моніторингу, персоналізованій аналітиці стану здоров'я та розробленні інтелектуальних систем підтримки прийняття рішень. До таких показників належать частота серцевих скорочень, частота дихання, насичення крові киснем, артеріальний тиск, температура тіла, показники варіабельності серцевого ритму та інші фізіологічні характеристики. Серед них частота серцевих скорочень є одним із найпоширеніших, найзрозуміліших і найважливіших показників, оскільки вона швидко реагує на зміни фізичного навантаження, стресу, болю, медикаментозного впливу, патологічних станів і загального функціонального стану організму.

У клінічній практиці частота серцевих скорочень традиційно визначається за електрокардіографічним сигналом або за даними медичних моніторів. Водночас у багатьох прикладних задачах дедалі більшого значення набувають неінвазивні та зручні методи реєстрації фізіологічних сигналів [5]. Одним із таких методів є фотоплетизмографія, яка ґрунтується на оптичній реєстрації змін кровонаповнення тканин. PPG-сигнал широко використовується у пульсоксиметрах, смартгодинниках, фітнес-браслетах, медичних моніторах та інших пристроях, що здійснюють безперервний або періодичний контроль стану людини.

Задача оцінювання частоти серцевих скорочень за PPG-сигналом на перший погляд може здаватися простою: кожне серцеве скорочення утворює пульсову хвилю, а тому достатньо знайти послідовні піки сигналу та визначити інтервали між ними. Однак у реальних клінічних і побутових умовах PPG-сигнал часто містить шум, пропуски, артефакти руху, зміну амплітуди, спотворення форми хвилі, ефекти недостатнього контакту датчика та апаратні особливості реєстрації. Через це класичні алгоритми пошуку піків можуть працювати нестабільно: на чистих фрагментах вони дають дуже точний результат, але на складних ділянках можуть створювати значні помилки.

Актуальність роботи зумовлена потребою в розробленні більш стійких методів оцінювання біометричних показників людини за фізіологічними сигналами. Методи машинного та глибокого навчання дають змогу не обмежуватися ручним пошуком піків, а використовувати ширшу інформацію про форму, ритмічність і локальні властивості сигналу [9-12]. Особливо перспективними є компактні нейронні моделі для одновимірних часових рядів, оскільки вони можуть працювати без надмірного обсягу параметрів і водночас враховувати структуру сигналу, яку складно повністю описати набором ручних ознак.

Тема кваліфікаційної роботи – «Оцінка біометричних показників людини з використанням глибокого навчання». У межах цієї теми як цільовий біометричний показник обрано частоту серцевих скорочень, а як вхідний фізіологічний сигнал – фотоплетизмографічний сигнал. Такий вибір є обґрунтованим, оскільки ЧСС є базовим біометричним показником стану людини, а PPG є поширеним, доступним і практично важливим способом його непрямого оцінювання.

Об'єктом дослідження є фізіологічні часові ряди, що використовуються для оцінювання біометричних показників людини.

Предметом дослідження є методи попередньої обробки фотоплетизмографічного сигналу та моделі машинного і глибокого навчання для оцінювання частоти серцевих скорочень.

Метою кваліфікаційної роботи є підвищення точності та стійкості автоматизованого оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом порівняно з класичним статистичним методом шляхом створення програмного засобу на основі методів машинного та глибокого навчання. Досягнення мети перевіряється за показниками MAE, RMSE, R^2 , систематичного зсуву, процентилів абсолютної похибки та частки прогнозів у межах заданих порогів.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

1. Проаналізувати предметну область оцінювання біометричних показників людини за фізіологічними сигналами.
2. Розглянути особливості фотоплетизмографічного сигналу та його використання для оцінювання частоти серцевих скорочень.
3. Проаналізувати класичні методи оцінювання ЧСС за PPG-сигналом і сучасні підходи на основі машинного та глибокого навчання.
4. Вибрати джерело даних і релевантні канали фізіологічних сигналів для проведення експериментів.
5. Сформувати набір даних із 30-секундних PPG-вікон і відповідних референсних значень ЧСС.
6. Забезпечити коректне розбиття даних на навчальну, валідаційну й тестову вибірки без витоку між клінічними кейсами.
7. Реалізувати попередню обробку PPG-сигналу, зокрема заповнення пропусків, фільтрацію, зменшення частоти дискретизації та нормалізацію.
8. Реалізувати статистичний базовий метод оцінювання ЧСС на основі пошуку піків PPG-сигналу.
9. Побудувати модель XGBoost на основі ручних ознак, що описують статистичні, часові, спектральні та якісні характеристики PPG-вікна.
10. Розробити й навчити компактну нейронну модель Conv+LSTM.
11. Розробити й навчити компактну нейронну модель Conv+Transformer.
12. Провести порівняльне тестування моделей за метриками MAE, RMSE, R^2 , систематичного зсуву, процентилів абсолютної похибки та частки прогнозів у межах заданих порогів.
13. Проаналізувати результати на рівні окремих клінічних кейсів.
14. Дослідити залежність похибки від діапазону ЧСС і якості PPG-сигналу.
15. Оцінити компактність і швидкодію моделей.
16. Сформулювати висновки щодо доцільності використання компактних нейронних моделей для оцінювання ЧСС за PPG-сигналом.

У роботі використано відкритий набір клінічних даних VitalDB. Для кожного клінічного кейсу використовувався PPG-канал SNUADC/PLETH, а як референсна ЧСС – канал Solar8000/HR. Було сформовано набір даних із 6157 клінічних кейсів і 2129005 вікон тривалістю 30 секунд. Розбиття даних виконувалося за ідентифікатором клінічного кейсу, що дало змогу уникнути ситуації, коли схожі фрагменти одного запису потрапляють одночасно до навчальної та тестової вибірок.

Методи дослідження включають цифрову обробку сигналів, статистичне оцінювання за міжпіковими інтервалами, градієнтний бустинг дерев рішень, нейронні мережі для часових рядів, регресійний аналіз, оцінювання похибок, бутстреп із кластеризацією за клінічними кейсами, аналіз Бленда-Альтмана та аналіз залежності похибки від властивостей даних.

Науково-практичне значення роботи полягає в тому, що в ній побудовано не лише окрему модель, а повний відтворюваний підхід до оцінювання біометричного показника людини за PPG-сигналом. Запропонований підхід охоплює формування набору даних, попередню обробку, навчання кількох типів моделей, парне тестове порівняння, аналіз стабільності результатів на рівні клінічних кейсів і оцінювання практичної придатності моделей з погляду швидкодії та компактності.

Практичне значення роботи полягає у можливості використання отриманого програмного підходу як основи для подальших досліджень у сфері аналізу фізіологічних сигналів, розроблення систем моніторингу стану людини, побудови моделей для носимих пристроїв або медичних інформаційних систем. Отримані результати показують, що компактна модель Conv+LSTM може забезпечити високу точність оцінювання ЧСС у центральному діапазоні розподілу даних і має достатню швидкодію для практичного застосування.

Кваліфікаційна робота складається зі вступу, трьох розділів, висновків, переліку джерел посилання та одного додатка. У першому розділі розглянуто існуючі методи та підходи до оцінювання біометричних показників за фізіологічними сигналами. У другому розділі наведено постановку задачі,

методику формування даних, алгоритми попередньої обробки, опис моделей і програмної реалізації. У третьому розділі подано результати експериментального оцінювання, порівняння моделей, аналіз похибок, стабільності, обмежень і швидкодії.

РОЗДІЛ 1. ОГЛЯД І АНАЛІЗ МЕТОДІВ ОЦІНЮВАННЯ БІОМЕТРИЧНИХ ПОКАЗНИКІВ ЛЮДИНИ ЗА ФІЗІОЛОГІЧНИМИ СИГНАЛАМИ

1.1. Біометричні показники людини як об'єкт комп'ютерного аналізу

Біометричні показники людини є кількісними або якісними характеристиками організму, які можуть бути використані для оцінювання фізіологічного стану, ідентифікації, моніторингу здоров'я, прогнозування ризиків або підтримки прийняття рішень. У широкому розумінні до біометричних показників можуть належати як сталі або повільно змінні характеристики людини, наприклад риси обличчя чи відбитки пальців, так і динамічні фізіологічні показники, що змінюються в часі. У цій роботі увага зосереджена саме на динамічних фізіологічних показниках, оскільки вони відображають поточний стан організму та можуть змінюватися під впливом навантаження, лікування, болю, стресу, анестезії, патологічних станів і зовнішніх умов.

Фізіологічні часові ряди є важливим джерелом інформації для комп'ютерного аналізу. До них належать електрокардіографічний сигнал, фотоплетизмографічний сигнал, сигнали дихання, електроенцефалографія, артеріальний тиск, насичення крові киснем, температура та інші показники. На відміну від статичних ознак, фізіологічні сигнали мають часову структуру. Це означає, що для їх аналізу важливі не лише окремі значення, а й форма сигналу, ритм, періодичність, локальні зміни, тренди, повторюваність і взаємозв'язок між сусідніми фрагментами.

Комп'ютерний аналіз біометричних показників може виконувати кілька типів задач. Перший тип – це пряме вимірювання або оцінювання фізіологічного параметра за сигналом. До нього належить задача оцінювання частоти серцевих скорочень за PPG-сигналом. Другий тип – класифікація станів, наприклад виявлення патологічних ритмів, епізодів апное або артефактів. Третій тип – прогнозування майбутніх значень або ризиків, наприклад прогнозування

погіршення стану пацієнта. Четвертий тип – оцінювання якості сигналу, тобто визначення того, чи можна довіряти отриманому вимірюванню.

З погляду комп'ютерних наук фізіологічні сигнали є складними об'єктами аналізу з кількох причин. По-перше, вони мають шумову природу. Реєстрація сигналів залежить від датчика, середовища, рухів людини, якості контакту та апаратних параметрів. По-друге, сигнали мають індивідуальну варіативність: одна й та сама фізіологічна характеристика може проявлятися по-різному в різних людей. По-третє, у клінічних даних часто трапляються пропуски, артефакти та нетипові значення. По-четверте, розподіл фізіологічних станів зазвичай є нерівномірним: нормальні або типові стани представлені значно краще, ніж рідкісні або критичні.

Через ці особливості розроблення моделей для аналізу біометричних показників потребує поєднання методів цифрової обробки сигналів, статистичного аналізу, машинного навчання, глибокого навчання й ретельної експериментальної перевірки. Недостатньо навчити модель і оцінити її лише за однією середньою метрикою. Необхідно також аналізувати стабільність результатів на рівні окремих записів, поведінку моделі в різних діапазонах цільового показника, вплив якості сигналу, наявність систематичного зсуву та обмеження моделі.

У цій роботі як біометричний показник розглядається частота серцевих скорочень. Її вибір зумовлений тим, що цей показник є фундаментальним для оцінювання стану людини, широко використовується в медичних і споживчих пристроях та може бути отриманий із PPG-сигналу. Водночас практичне оцінювання ЧСС за PPG залишається нетривіальною задачею через шум, артефакти, варіативність форми пульсової хвилі та неоднозначність референсних значень у реальних клінічних даних.

1.2. Частота серцевих скорочень як біометричний і фізіологічний показник

Частота серцевих скорочень є кількістю скорочень серця за одиницю часу, зазвичай за одну хвилину. Вона є одним із найпоширеніших фізіологічних показників, оскільки прямо пов'язана з роботою серцево-судинної системи та швидко реагує на зміну стану організму. У клінічному середовищі ЧСС використовується під час моніторингу пацієнтів у палатах інтенсивної терапії, операційних, післяопераційних відділеннях, а також у загальній діагностиці й спостереженні.

Частота серцевих скорочень є важливою не лише як окреме числове значення. Її зміни в часі можуть відображати реакцію організму на навантаження, біль, медикаменти, крововтрату, порушення дихання, зміну рівня кисню, емоційний стан або розвиток патологічних процесів. У задачах цифрового моніторингу ЧСС може використовуватися як самостійний показник або як складова більш складних індексів і моделей ризику.

З погляду обробки сигналів ЧСС може оцінюватися різними способами. Найбільш прямий шлях – використання електрокардіографічного сигналу, де серцеві скорочення проявляються у вигляді R-піків. Інший поширений шлях – використання PPG-сигналу, у якому серцеві скорочення проявляються через пульсові хвилі. Також ЧСС може бути доступною як готовий моніторний показник у клінічних системах.

Електрокардіографія зазвичай вважається більш прямим джерелом інформації про електричну активність серця. Проте PPG має свої переваги: його можна отримувати оптичним датчиком, він широко використовується у пульсоксиметрах і носимих пристроях, а його реєстрація часто є простішою з погляду користувача [7, 8]. Саме тому задача оцінювання ЧСС за PPG має велику практичну значущість.

Однак між електричною активністю серця та PPG-хвилею існує фізіологічний і технічний проміжок. PPG відображає не саме електричне скорочення серця, а зміну кровонаповнення тканин після проходження пульсової

хвилі [5]. Тому PPG-сигнал може затримуватися відносно електрокардіографічного сигналу, змінювати форму залежно від судинного тону, перфузії та місця вимірювання. Це створює додаткові труднощі для точного оцінювання ЧСС.

У контексті цієї роботи ЧСС розглядається як цільова змінна регресійної задачі. Для кожного 30-секундного фрагмента PPG потрібно отримати одне числове значення ЧСС в ударах за хвилину. Така постановка не вимагає від моделі точного визначення кожного окремого серцевого циклу, але потребує правильного узагальнення ритмічної інформації протягом усього вікна.

1.3. Фотоплетизмографічний сигнал: природа, структура та особливості

Фотоплетизмографія є методом оптичної реєстрації змін об'єму крові в тканинах [5, 6]. Типовий PPG-датчик містить джерело світла та фотоприймач. Світло проходить через тканину або відбивається від неї, а фотоприймач фіксує зміну інтенсивності світла, пов'язану зі змінами кровонаповнення. Оскільки кров поглинає світло, пульсові зміни об'єму крові створюють характерні коливання сигналу.

PPG-сигнал має кілька компонентів. Повільна складова пов'язана з базовим рівнем поглинання, властивостями тканин, положенням датчика, рухами та іншими довготривалими факторами. Пульсова складова пов'язана зі змінами кровонаповнення внаслідок серцевих скорочень. Для оцінювання ЧСС найбільший інтерес становить саме пульсова складова, оскільки її періодичність відповідає серцевому ритму.

Типова пульсова хвиля PPG має підйом, пік і спад. У деяких випадках на спаді може бути помітна додаткова хвиля або перегин, пов'язаний із відбиттям пульсової хвилі в судинній системі. Проте в реальних даних форма хвилі може суттєво змінюватися. На неї впливають артеріальний тиск, судинний тонус, перфузія, рухи, якість контакту датчика, технічні параметри реєстрації й стан пацієнта.

Для оцінювання ЧСС можна використовувати кілька властивостей PPG-сигналу. Найпростіша властивість – інтервали між піками. Якщо піки виявлені правильно, середній або медіанний міжпіковий інтервал дозволяє оцінити частоту серцевих скорочень. Однак якщо піки нечіткі, подвійні, пропущені або спотворені шумом, цей підхід може давати великі помилки.

Інша група властивостей пов'язана з частотним складом сигналу. Оскільки серцеві скорочення мають певну частоту, у спектрі PPG-сигналу може бути помітний домінантний компонент, пов'язаний із пульсом. Проте частотний аналіз також має обмеження, особливо на коротких вікнах, у разі зміни ЧСС у межах вікна або за наявності шуму.

Сучасні методи машинного та глибокого навчання можуть використовувати ширший набір ознак. Наприклад, згорткові нейронні мережі можуть автоматично виділяти локальні патерни форми хвилі, а рекурентні моделі або моделі з механізмом уваги можуть агрегувати часову структуру протягом усього вікна. Це дозволяє моделі не залежати повністю від одного етапу пошуку піків.

У цій роботі PPG-сигнал попередньо фільтрується, зменшується за частотою дискретизації та нормалізується. Така обробка потрібна для того, щоб зменшити вплив повільного дрейфу, високочастотного шуму та різного масштабу амплітуди. Водночас обробка не повинна знищити ритмічну та морфологічну інформацію, необхідну для оцінювання ЧСС.

1.4. Класичні методи оцінювання частоти серцевих скорочень за PPG-сигналом

Класичні методи оцінювання ЧСС за PPG-сигналом ґрунтуються переважно на цифровій обробці сигналів [7, 8]. Найпоширенішою ідеєю є попередня фільтрація сигналу, пошук пульсових піків і розрахунок міжпікових інтервалів. Якщо між двома послідовними піками проходить певний час, то частота серцевих скорочень може бути оцінена як 60, поділене на тривалість цього інтервалу в секундах.

Перевагою такого підходу є інтерпретованість. Кожен етап можна пояснити: фільтрація прибирає шум, пошук піків знаходить пульсові хвилі, міжпікові інтервали відображають серцеві цикли. Крім того, такі методи зазвичай є швидкими й не потребують навчальних даних. Це робить їх зручними для простих систем і базових порівнянь.

Проте класичний підхід має суттєві недоліки. Він сильно залежить від якості сигналу. Якщо PPG-хвиля добре виражена, піки знаходяться правильно, а оцінка ЧСС є точною. Якщо сигнал шумний, має рухові артефакти, слабку амплітуду або деформовану форму, алгоритм може знаходити зайві піки або пропускати справжні. У такому разі навіть одна помилка у визначенні піку може суттєво змінити оцінку ЧСС.

Ще одна складність полягає у виборі параметрів пошуку піків. Необхідно задавати мінімальну відстань між піками, мінімальну виразність піку, діапазон допустимих міжпікових інтервалів та інші параметри. Якщо параметри підібрані занадто жорстко, алгоритм може пропускати правильні піки. Якщо занадто м'яко – знаходити багато хибних піків. Універсальний набір параметрів, який однаково добре працює для всіх пацієнтів і всіх умов, підібрати складно.

Для зменшення впливу окремих помилок часто використовують медіанний міжпіковий інтервал, а не середній. Медіана є стійкішою до одиничних викидів. Також застосовують фізіологічні обмеження: наприклад, відкидати інтервали, які відповідають занадто низькій або занадто високій ЧСС. У цій роботі статистичний метод використовує саме таку логіку: PPG-вікно центрується, у ньому шукаються піки, міжпікові інтервали фільтруються за діапазоном 40–180 уд/хв, а фінальний прогноз визначається через медіанний інтервал.

Попри обмеження, класичний статистичний підхід є важливим для цієї роботи. Він виконує роль базового методу, з яким порівнюються навчальні моделі. Якщо нейронна модель не перевершує простий інтерпретований підхід, її практична доцільність є сумнівною. Якщо ж нейронна модель показує суттєво кращу стабільність і менший хвіст великих помилок, це є аргументом на користь використання глибокого навчання.

У проведених експериментах статистичний підхід справді показав подвійну поведінку. Він мав низьку медіанну абсолютну похибку, тобто часто був дуже точним. Водночас він мав великий 95-й перцентиль абсолютної похибки та найгірше значення RMSE. Це підтверджує типову проблему пікового підходу: він добре працює на чистих сигналах, але нестабільний на складних вікнах.

1.5. Методи машинного навчання в задачах аналізу фізіологічних сигналів

Методи машинного навчання дають змогу перейти від жорстко заданих правил до моделей, які навчаються на даних [10, 11]. У задачах аналізу фізіологічних сигналів це особливо важливо, оскільки реальні сигнали мають високу варіативність, а їхня форма не завжди добре описується простими правилами. Замість того щоб вручну задавати всі можливі випадки, можна побудувати модель, яка вчиться зіставляти властивості сигналу з цільовим фізіологічним показником.

Одним із поширених підходів є перетворення сигналу на набір ручних ознак. У такому випадку часовий ряд спочатку стискається до вектора числових характеристик. Це можуть бути статистичні ознаки, перцентилі, ознаки похідної, спектральні ознаки, показники якості сигналу, ознаки періодичності або локальні характеристики в окремих сегментах вікна. Після цього до отриманого вектора застосовується модель машинного навчання.

Перевагою такого підходу є відносна простота. Табличні моделі зазвичай швидко навчаються, не потребують дуже складної інфраструктури та можуть працювати з помірною кількістю ознак. Крім того, ручні ознаки часто мають зрозумілу інтерпретацію. Наприклад, стандартне відхилення сигналу може характеризувати його варіативність, спектральні ознаки – частотну структуру, а показники якості – надійність вікна.

Серед моделей машинного навчання для табличних ознак широко застосовуються дерева рішень, випадкові ліси, градієнтний бустинг, лінійні

моделі та методи опорних векторів. У цій роботі як сильний табличний базовий підхід використано XGBoost [12]. Його вибір обґрунтований тим, що градієнтний бустинг дерев рішень добре працює з нелінійними залежностями, може використовувати різноманітні ознаки та часто демонструє високу якість на табличних задачах.

У межах роботи для XGBoost сформовано 133 ознаки. Вони описують якість PPG-вікна, глобальну статистику сигналу, процентилі, локальні середні значення та стандартні відхилення в сегментах, а також спектральні характеристики. Такий набір ознак дає моделі інформацію про загальний масштаб, форму, локальну зміну та частотний склад сигналу.

Однак підхід із ручними ознаками має обмеження. Перетворення сигналу на фіксований вектор неминуче призводить до втрати частини часової інформації. Наприклад, дві послідовності можуть мати схожі середні значення, стандартні відхилення й навіть подібні спектральні ознаки, але різну локальну форму хвилі. Для задачі оцінювання ЧСС така форма може бути важливою, особливо якщо сигнал містить артефакти або нерівномірний ритм.

Ще одним обмеженням є те, що якість табличної моделі залежить від якості ручних ознак. Якщо інженер ознак не передбачив певну важливу характеристику сигналу, модель не зможе її використати. Натомість нейронні мережі можуть автоматично формувати внутрішні ознаки з повного сигналу. Саме тому у цій роботі XGBoost розглядається як сильний, але не фінальний підхід: він дозволяє перевірити, наскільки далеко можна зайти з ручними ознаками, і порівняти це з моделями глибокого навчання.

1.6. Глибоке навчання для обробки одновимірних часових рядів

Глибоке навчання є підходом, у якому модель автоматично формує багаторівневі представлення даних [7]. У задачах аналізу зображень це проявляється у виділенні локальних візуальних ознак, а в задачах часових рядів – у виділенні локальних і довготривалих часових закономірностей. Для PPG-

сигналу це особливо важливо, оскільки інформація про ЧСС міститься не лише в окремих піках, а й у ритмічній структурі всього фрагмента сигналу.

Для одновимірних фізіологічних сигналів часто використовують згорткові нейронні мережі. Одновимірні згортки дозволяють знаходити локальні патерни в часовому ряді: підйоми, спади, піки, різкі зміни, локальні коливання. Згорткові шари можуть поступово зменшувати довжину сигналу й перетворювати сирі відліки на послідовність векторних ознак. Це робить подальшу обробку ефективнішою.

Рекурентні нейронні мережі, зокрема LSTM, призначені для роботи з послідовностями [13]. Вони дозволяють враховувати залежності між елементами послідовності та накопичувати інформацію про попередні стани. Для PPG-сигналу це означає можливість аналізувати не лише локальну форму хвилі, а й загальний ритм протягом усього вікна. Двонапрямлена LSTM додатково враховує послідовність в обох напрямках, що корисно для задачі, де прогноз формується після перегляду всього вікна.

Моделі Transformer використовують механізм уваги. На відміну від рекурентних мереж, вони можуть прямо моделювати взаємозв'язки між різними позиціями послідовності [14]. У задачах обробки тексту й багатьох типів послідовностей Transformer показав високу ефективність. Для фізіологічних сигналів його застосування також є перспективним, але потребує правильного перетворення довгого сигналу на послідовність ознак, оскільки пряме використання дуже довгих послідовностей може бути обчислювально дорогим.

У цій роботі використано дві компактні нейронні архітектури. Перша – Conv+LSTM. Вона поєднує згортковий блок для виділення локальних ознак і двонапрямлену LSTM для узагальнення ритмічної часової структури. Друга – Conv+Transformer. Вона використовує згорткове перетворення сигналу на послідовність ознак, позиційне представлення та кодувальник Transformer.

Обидві моделі є компактними: вони мають близько 60 тисяч навчальних параметрів. Це важливо, оскільки метою роботи є не лише досягнення високої точності, а й перевірка того, чи можуть відносно невеликі моделі ефективно

оцінювати ЧСС за PPG-сигналом. Компактність моделі має практичне значення для потенційного використання в системах, де обмежені пам'ять, обчислювальні ресурси або час отримання прогнозу.

Порівняння Conv+LSTM і Conv+Transformer також має дослідницьку цінність. Воно дозволяє оцінити, який спосіб агрегації часової інформації краще працює для конкретної задачі: рекурентний механізм чи механізм уваги. У проведених експериментах Conv+LSTM показала кращу середню точність і стабільність, тоді як Conv+Transformer була швидшою під час прямого проходу, але мала більшу похибку та помітніший від'ємний зсув.

1.7. Набори даних для дослідження фізіологічних сигналів

Якість дослідження моделей машинного та глибокого навчання значною мірою залежить від якості й обсягу даних. Для задач аналізу фізіологічних сигналів важливо мати записи, які містять не лише окремі короткі приклади, а довгі часові ряди, різних пацієнтів, різні стани, реальні шуми та референсні показники. Саме такі дані дозволяють оцінити не лише лабораторну точність, а й поведінку моделі в умовах, наближених до реального застосування.

У роботі використано VitalDB – відкриту базу клінічних фізіологічних сигналів, описану в [6]. Цей набір даних містить клінічні записи, у яких доступні хвильові сигнали та моніторні фізіологічні показники. Для поставленої задачі важливо, що в VitalDB можна отримати фотоплетизмографічний сигнал і відповідний моніторний показник частоти серцевих скорочень. Це дає можливість сформулювати задачу контрольованого навчання, де вхідним об'єктом є PPG-вікно, а цільовою змінною – референсна ЧСС.

Використання реальних клінічних даних має суттєві переваги. По-перше, вони містять природну варіативність сигналів. По-друге, вони включають різні рівні якості PPG. По-третє, вони дозволяють оцінити моделі на великій кількості клінічних кейсів. У цій роботі використано 6157 кейсів, що дозволило сформувати понад два мільйони 30-секундних вікон.

Однак реальні клінічні дані мають і недоліки. Референсні канали не завжди є ідеальними. Моніторні показники можуть містити артефакти, пропуски або фізіологічно неправдоподібні значення. У тестовій вибірці цієї роботи було виявлено значення референсної ЧСС 0 і 300 уд/хв. Такі значення можуть суттєво впливати на метрики, особливо на RMSE, оскільки ця метрика сильно штрафує великі похибки.

Ще однією проблемою є нерівномірність розподілу ЧСС. Основна маса тестових вікон знаходиться в діапазоні приблизно 50–100 уд/хв. Високі та екстремальні значення представлені значно слабше. Це означає, що модель може добре працювати в основному діапазоні, але мати труднощі з рідкісними станами. Тому в роботі окремо проведено аналіз похибки за діапазонами ЧСС.

Під час роботи з такими даними особливу увагу потрібно приділяти способу розбиття на навчальну й тестову вибірки. Якщо розбивати дані випадково за вікнами, у навчальну й тестову вибірки можуть потрапити фрагменти одного й того самого клінічного запису. Це створює ризик витоку даних і завищення якості. У цій роботі розбиття виконано за ідентифікатором клінічного кейсу, тобто всі вікна одного клінічного кейсу належать лише одній вибірці.

1.8. Проблеми оцінювання якості моделей у задачах аналізу медичних часових рядів

Оцінювання якості моделей у задачах аналізу медичних часових рядів має кілька особливостей. На відміну від багатьох стандартних задач машинного навчання, тут окремі приклади часто не є незалежними. Наприклад, якщо довгий клінічний запис розбивається на багато 30-секундних вікон, сусідні вікна одного кейсу можуть мати схожу форму сигналу, схожу ЧСС і однакові апаратні умови реєстрації. Тому проста оцінка на рівні всіх вікон може створювати враження більшої статистичної надійності, ніж є насправді.

Через це важливо оцінювати моделі не лише на рівні окремих вікон, а й на рівні клінічних кейсів. Аналіз на рівні клінічних кейсів відповідає практичному

питанню: наскільки добре модель працює на новому записі пацієнта. Якщо модель має низьку середню похибку за всіма вікнами, але погано працює на значній частині кейсів, її практична цінність обмежена.

Інша проблема – вибір метрик. Середня абсолютна похибка є зрозумілою і зручною, але вона не повністю описує розподіл помилок. Корінь середньоквадратичної похибки краще виявляє великі помилки, але може бути надмірно чутливим до артефактів у референсі. Медіанна абсолютна похибка показує типову помилку, але може приховувати довгий хвіст великих похибок. Тому в роботі використано комплекс метрик.

Ще одна важлива характеристика – систематичний зсув. Модель може мати прийнятну середню абсолютну похибку, але систематично занижувати або завищувати ЧСС. У медичних задачах це важливо, оскільки постійне заниження високої ЧСС може бути небезпечним у практичному використанні. Саме тому в роботі аналізується не лише абсолютна похибка, а й середнє значення різниці між прогнозом і референсом.

Для оцінювання узгодженості прогнозів із референсом використовується аналіз Бленда-Альтмана [15]. Він дозволяє побачити не лише середній зсув, а й межі, в яких зазвичай знаходиться різниця між прогнозом і референсом. Це особливо корисно для порівняння моделей у задачах, близьких до вимірювальних.

Важливим є також аналіз похибки за підгрупами. Якщо загальна якість моделі висока, це не означає, що вона однаково добре працює для всіх діапазонів ЧСС. У цій роботі окремо аналізуються низькі, центральні, високі та екстремальні діапазони. Також аналізується залежність похибки від якості PPG-сигналу. Це дозволяє визначити, у яких умовах модель є надійною, а в яких її результати потребують обережної інтерпретації.

1.9. Мотивований висновок щодо необхідності виконання роботи

Проведений огляд показує, що оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом є актуальною задачею на перетині обробки

сигналів, машинного навчання, глибокого навчання та аналізу медичних даних. Частота серцевих скорочень є важливим біометричним показником, а PPG-сигнал є поширеним і практично зручним джерелом інформації про пульсову хвилю.

Класичні методи на основі пошуку піків мають суттєву перевагу у вигляді простоти та інтерпретованості. Проте вони залежать від якості сигналу й можуть створювати великі помилки у складних вікнах. Методи машинного навчання на основі ручних ознак дозволяють враховувати ширший набір характеристик сигналу, але все одно залежать від того, наскільки повно ці ознаки описують часову структуру PPG. Методи глибокого навчання дають змогу працювати безпосередньо з формою сигналу й автоматично виділяти інформативні представлення.

Водночас використання глибокого навчання саме по собі не гарантує кращого результату. Необхідно коректно сформувати набір даних, уникнути витоку між клінічними кейсами, виконати якісну попередню обробку, порівняти нейронні моделі з сильними базовими підходами, оцінити результати не лише за середньою похибкою, а й за стабільністю, зсувом, поведінкою в різних діапазонах ЧСС і залежністю від якості сигналу.

Саме тому виконання цієї роботи є доцільним. Вона спрямована не лише на побудову однієї моделі, а на комплексне дослідження задачі оцінювання біометричного показника людини за PPG-сигналом. У роботі порівнюються статистичний підхід, таблична модель машинного навчання та дві компактні нейронні архітектури. Такий підхід дозволяє обґрунтовано визначити, чи дає глибоке навчання практичну перевагу в цій задачі, які саме обмеження залишаються та в яких напрямках доцільно продовжувати дослідження.

1.10. Висновки до розділу 1

У першому розділі було розглянуто біометричні показники людини як об'єкт комп'ютерного аналізу та визначено місце частоти серцевих скорочень серед динамічних фізіологічних показників. Показано, що ЧСС є одним із

базових індикаторів стану людини, який широко використовується в клінічному моніторингу, носимих пристроях, пульсоксиметрах, системах дистанційного спостереження та медичних інформаційних системах.

Проаналізовано фотоплетизмографічний сигнал як джерело інформації для оцінювання ЧСС. Встановлено, що PPG-сигнал є практично зручним і поширеним способом реєстрації пульсової хвилі, однак його обробка ускладнюється шумом, артефактами руху, зміною амплітуди, пропусками, варіативністю форми сигналу та залежністю від умов реєстрації.

Розглянуто класичні методи оцінювання ЧСС за PPG-сигналом, методи машинного навчання на основі ручних ознак і підходи глибокого навчання для одновимірних часових рядів. Показано, що класичні пікові методи є простими й інтерпретованими, але чутливими до якості сигналу, тоді як моделі машинного та глибокого навчання можуть використовувати ширшу інформацію про статистичні, часові, спектральні та морфологічні властивості PPG-вікна.

Окремо розглянуто питання вибору набору даних і оцінювання якості моделей у задачах аналізу медичних часових рядів. Обґрунтовано необхідність розбиття даних за клінічними кейсами, використання кількох метрик якості, аналізу систематичного зсуву, стабільності результатів на рівні клінічних записів і поведінки моделей у різних діапазонах ЧСС.

Отже, у розділі обґрунтовано актуальність задачі автоматизованого оцінювання частоти серцевих скорочень за PPG-сигналом і показано доцільність використання методів машинного та глибокого навчання для створення програмного засобу, який має забезпечити підвищення точності та стійкості оцінювання порівняно з класичним статистичним методом.

РОЗДІЛ 2. ПОСТАНОВКА ЗАДАЧІ, МЕТОДИКА ТА ПРОГРАМНА РЕАЛІЗАЦІЯ ОЦІНЮВАННЯ ЧАСТОТИ СЕРЦЕВИХ СКОРОЧЕНЬ ЗА PPG-СИГНАЛОМ

2.1. Постановка задачі дослідження

У межах кваліфікаційної роботи розв'язується задача оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом. З математичного погляду ця задача формулюється як задача регресії за часовим рядом. На вхід подається фрагмент PPG-сигналу фіксованої тривалості, а на виході потрібно отримати числове значення ЧСС в ударах за хвилину.

Нехай x_i – 30-секундний фрагмент PPG-сигналу, а y_i – відповідне референсне значення ЧСС. Тоді потрібно побудувати функцію:

$$\hat{y}_i = f(x_i), \quad (2.1)$$

де x_i – 30-секундний фрагмент PPG-сигналу;

y_i – референсне значення ЧСС;

$\hat{y}_i$ – прогнозоване моделлю значення ЧСС;

f – функція оцінювання ЧСС.

У статистичному підході ця функція задається вручну через пошук піків і міжпікові інтервали. У підходах машинного та глибокого навчання функція f навчається на даних.

Вхідний сигнал у роботі має дві форми. Перша – сирий PPG-сигнал із частотою дискретизації 500 Гц. Для 30-секундного вікна це 15000 відліків. Друга – підготовлений сигнал із частотою 125 Гц після фільтрації, зменшення частоти дискретизації та нормалізації. Для 30-секундного вікна це 3750 відліків. Саме підготовлений сигнал використовується для статистичного підходу та нейронних моделей.

Цільова змінна формується на основі каналу Solar8000/HR. Для кожного 30-секундного вікна визначається референсне значення ЧСС. У роботі

воно розглядається як значення, з яким порівнюються прогнози моделей. Водночас важливо враховувати, що цей канал не є ідеальним еталоном і може містити аномальні значення.

Основна складність задачі полягає в тому, що PPG-сигнал є шумним фізіологічним часовим рядом. Не всі вікна мають однакову якість, не всі діапазони ЧСС представлені однаково, а референсні значення можуть містити артефакти. Тому задача не зводиться до простого навчання регресійної моделі. Потрібно побудувати повний методичний підхід, який включає формування даних, контроль якості, коректне розбиття, навчання кількох моделей і детальний аналіз результатів.

2.2. Об'єкт, предмет, мета та завдання роботи

Об'єктом дослідження є фізіологічні часові ряди, що відображають біометричні показники людини.

Предметом дослідження є методи попередньої обробки фотоплетизмографічного сигналу та моделі машинного і глибокого навчання для оцінювання частоти серцевих скорочень.

Метою кваліфікаційної роботи є підвищення точності та стійкості автоматизованого оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом порівняно з класичним статистичним методом, що перевіряється за показниками MAE, RMSE, R^2 , систематичного зсуву, P95 абсолютної похибки та частки прогнозів із похибкою не більше 5 і 10 уд/хв.

Для досягнення мети необхідно виконати такі завдання:

1. Проаналізувати існуючі підходи до оцінювання ЧСС за фізіологічними сигналами.
2. Обґрунтувати використання PPG-сигналу як джерела інформації для оцінювання ЧСС.
3. Вибрати набір даних і канали сигналів для експериментального дослідження.

4. Сформувати набір 30-секундних PPG-вікон із відповідними референсними значеннями ЧСС.
5. Реалізувати розбиття даних за клінічними кейсами.
6. Розробити алгоритм попередньої обробки PPG-сигналу.
7. Розрахувати показники якості PPG-вікон.
8. Реалізувати статистичний метод оцінювання ЧСС за піками PPG.
9. Побудувати модель XGBoost на основі ручних ознак.
10. Розробити модель Conv+LSTM для роботи з повним PPG-вікном.
11. Розробити модель Conv+Transformer для роботи з повним PPG-вікном.
12. Провести навчання й тестування моделей.
13. Порівняти моделі за основними регресійними метриками.
14. Оцінити стабільність моделей на рівні клінічних кейсів.
15. Дослідити вплив діапазону ЧСС і якості PPG-сигналу на похибку.
16. Проаналізувати швидкодію та компактність моделей.
17. Сформулювати обмеження роботи й напрями подальшого розвитку.

2.3. Обмеження задачі та припущення дослідження

Під час постановки задачі було прийнято кілька обмежень і припущень, які визначають межі застосування отриманих результатів.

По-перше, у роботі використовується один основний вхідний сигнал – PPG. Інші фізіологічні канали, наприклад електрокардіографія або артеріальний тиск, не використовуються як вхідні ознаки фінальних моделей. Це зроблено для того, щоб дослідити саме можливість оцінювання ЧСС за фотоплетизмографічним сигналом.

По-друге, референсною цільовою змінною є канал Solar8000/HR. Він розглядається як доступний клінічний показник ЧСС, але не як абсолютно безпомилковий еталон. У даних було виявлено екстремальні значення, зокрема 0 і 300 уд/хв. Тому результати потрібно інтерпретувати з урахуванням можливої недосконалості референсу.

По-третє, модель оцінює середнє значення ЧСС для 30-секундного вікна. Вона не призначена для визначення точного часу кожного окремого серцевого скорочення. Це відрізняє задачу від задачі детекції піків або аналізу варіабельності серцевого ритму.

По-четверте, оцінювання виконується на даних VitalDB. Хоча тестова вибірка містить клінічні кейси, які не використовувалися під час навчання, це все ще внутрішня перевірка в межах одного джерела даних. Для тверджень про загальну клінічну застосовність потрібна зовнішня перевірка на незалежних наборах даних.

По-п'яте, основна увага приділяється порівнянню компактних моделей. У роботі не ставиться задача побудувати максимально велику нейронну мережу. Навпаки, досліджується, чи можуть відносно невеликі архітектури забезпечити високу якість при помірних обчислювальних витратах.

По-шосте, навчання та тестування моделей виконувалося в локальному обчислювальному середовищі на базі процесора Apple Silicon M5 Pro. Цей факт не є предметом дослідження, але він характеризує експериментальне середовище й підтверджує, що отримані результати були досягнуті без використання великої серверної інфраструктури. У тексті роботи це зазначається лише як частина опису умов експериментів.

2.4. Загальна декомпозиція програмного рішення

Розроблене програмне рішення можна поділити на кілька взаємопов'язаних підсистем:

- підсистема завантаження клінічних сигналів;
- підсистема формування набору даних;
- підсистема попередньої обробки PPG;
- підсистема розрахунку показників якості;
- підсистема побудови ручних ознак;
- підсистема навчання моделей;
- підсистема формування прогнозів;

- підсистема експериментального оцінювання;
- підсистема побудови таблиць і графіків.

Підсистема завантаження клінічних сигналів відповідає за отримання потрібних каналів VitalDB для кожного клінічного кейсу. Вона завантажує PPG-сигнал і референсний HR-канал, перевіряє їх наявність і передає дані до наступних етапів.

Підсистема формування набору даних перетворює довгі часові ряди на набір 30-секундних вікон. Для кожного вікна зберігаються часові межі, ідентифікатор кейсу, референсне значення ЧСС, показники якості та сам сигнал.

Підсистема попередньої обробки виконує заповнення пропусків, фільтрацію, зменшення частоти дискретизації та нормалізацію. Її результатом є підготовлений сигнал, який має однакову довжину для всіх вікон і є придатним для подальшого аналізу.

Підсистема формування прогнозів охоплює чотири підходи: статистичний метод на основі піків PPG-сигналу, XGBoost, Conv+LSTM та Conv+Transformer. Для статистичного методу навчання не потрібне. Для XGBoost формується вектор ручних ознак. Для нейронних моделей використовується повне підготовлене PPG-вікно.

Підсистема оцінювання об'єднує прогнози всіх моделей на спільній тестовій вибірці та розраховує метрики. Вона також виконує аналіз на рівні клінічних кейсів, бутстреп-оцінювання довірчих інтервалів, аналіз Бленда-Альтмана, аналіз за діапазонами ЧСС та якістю сигналу.

Як показано на рисунку 2.1, розроблене рішення охоплює повний цикл роботи з даними: завантаження клінічних сигналів, формування вікон, попередню обробку, навчання моделей, формування прогнозів та експериментальне оцінювання.

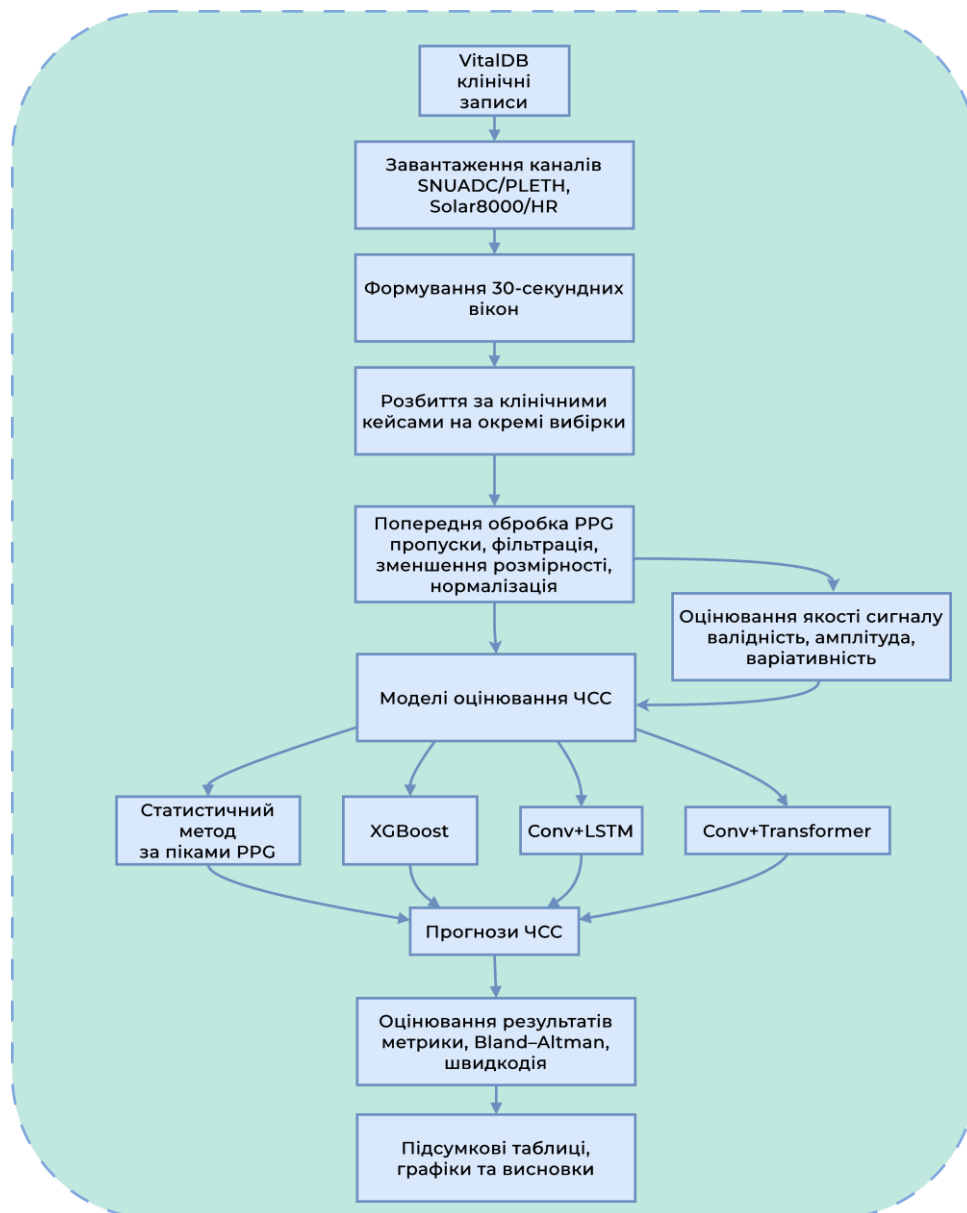


Рисунок 2.1 – Загальна схема програмного рішення для оцінювання ЧСС за PPG-сигналом

2.5. Вибір джерела даних і каналів фізіологічних сигналів

Для експериментального дослідження обрано VitalDB, оскільки цей набір даних містить клінічні записи з фізіологічними хвильовими сигналами та моніторними показниками. Його використання дозволяє досліджувати задачу на реальних клінічних даних, а не на штучно згенерованих або надто очищених лабораторних прикладах.

У межах роботи використано два основні канали:

- SNUADC/PLETH – фотоплетизмографічний сигнал;

- Solar8000/HR – референсна частота серцевих скорочень.

Канал SNUADC/PLETH обрано як вхідний сигнал, оскільки він відповідає PPG-сигналу, з якого можна оцінювати пульсову активність. Канал Solar8000/HR обрано як референсне значення ЧСС, оскільки він є готовим моніторним показником, доступним для великої кількості клінічних кейсів.

Додатково на ранньому етапі було виконано дослідження можливості отримання ЧСС з електрокардіографічного сигналу. Для цього використовувався канал SNUADC/ECG_II і алгоритми пошуку R-піків. Це дослідження допомогло перевірити логіку віконного аналізу та характер узгодження між різними джерелами ЧСС. Однак фінальні моделі навчалися саме на референсному каналі Solar8000/HR, оскільки він доступний для масштабного формування набору даних.

Вибір одного основного PPG-каналу як вхідного сигналу робить постановку задачі чіткою. Модель не використовує додаткову інформацію з інших каналів, тому результат відображає саме здатність обраних методів оцінювати ЧСС за фотоплетизмографією.

2.6. Формування набору даних із 30-секундних вікон

Довгі записи PPG-сигналу було розбито на неперекривні вікна тривалістю 30 секунд. Крок між вікнами також становив 30 секунд. Такий підхід означає, що кожен фрагмент сигналу використовується лише один раз, а сусідні вікна не перекриваються.

Вибір тривалості 30 секунд є компромісом між точністю й часовою роздільністю. За такої довжини навіть при низькій ЧСС у вікні є достатня кількість серцевих циклів для оцінювання ритму. Водночас вікно не є надто довгим, тому воно дозволяє відстежувати зміни ЧСС у межах клінічного запису.

Початкова частота дискретизації PPG становить 500 Гц. Отже, сире 30-секундне вікно містить $30 \times 500 = 15000$ відліків. Після попередньої обробки та зменшення частоти дискретизації до 125 Гц модельне вікно містить $30 \times 125 = 3750$ відліків.

Для кожного вікна зберігаються:

- назва вибірки;
- ідентифікатор клінічного кейсу;
- номер вікна;
- час початку вікна;
- час завершення вікна;
- середина вікна;
- референсне значення ЧСС;
- частка валідних PPG-відліків;
- міжквартильний розмах амплітуди PPG;
- стандартне відхилення підготовленого PPG;
- частка потенційно обмежених значень;
- сирий PPG-сигнал;
- підготовлений PPG-сигнал.

Такий формат дозволяє використовувати один і той самий набір даних для різних моделей. Статистичний метод PPG і нейронні моделі працюють із підготовленим сигналом. XGBoost використовує ручні ознаки, частина яких обчислюється з підготовленого сигналу, а частина – з показників якості.

Структуру одного запису сформованого набору даних наведено в таблиці 2.1.

Як видно з таблиці 2.1, один запис містить не лише цільове значення ЧСС, а й технічні поля, показники якості та два представлення PPG-сигналу – сире й підготовлене.

Таблиця 2.1 – Структура одного запису сформованого набору даних

Поле запису	Технічне позначення	Опис
Назва вибірки	split	Визначає, до якої частини набору даних належить вікно: навчальної, валідаційної або тестової.
Ідентифікатор клінічного кейсу	case_id	Використовується для групування всіх вікон одного клінічного запису та розбиття даних без витоку між вибірками.
Номер вікна в межах кейсу	window_id	Порядковий номер 30-секундного фрагмента у відповідному клінічному записі.
Час початку вікна	start_sec	Час початку фрагмента відносно початку клінічного запису, с.
Час завершення вікна	end_sec	Час завершення фрагмента відносно початку клінічного запису, с.
Середина вікна	window_mid_sec	Центральний момент часу 30-секундного фрагмента, с.
Референсна ЧСС	label_ref_hr_bpm	Цільове значення частоти серцевих скорочень для відповідного вікна, уд/хв.
Частка валідних відліків PPG-сигналу	ppg_valid_ratio	Показник повноти сирого PPG-вікна, що визначає частку скінчених і придатних для обробки відліків.
Міжквартильний розмах амплітуди PPG-сигналу	ppg_amplitude_iqr	Стійка характеристика варіативності амплітуди сирого PPG-сигналу.

Продовження таблиці 2.1

Поле запису	Технічне позначення	Опис
Стандартне відхилення підготовленого PPG- сигналу	ppg_prepared_std	Характеристика варіативності PPG-сигналу після заповнення пропусків, фільтрації та нормалізації.
Частка потенційно обмежених значень сигналу	ppg_clipped_ratio	Показник можливої наявності насичення або обрізання PPG-сигналу за амплітудою.
Частота дискретизації сирого сигналу	raw_fs_hz	Початкова частота дискретизації PPG-сигналу, Гц.
Частота дискретизації підготовленого сигналу	model_fs_hz	Частота дискретизації PPG-сигналу після зменшення частоти, Гц.
Кількість відліків сирого сигналу	raw_samples	Кількість відліків у сирому 30-секундному PPG-вікні.
Кількість відліків підготовленого сигналу	model_samples	Кількість відліків у підготовленому PPG-вікні після зменшення частоти дискретизації.
Сирий PPG-сигнал	ppg_raw	Початковий фрагмент фотоплетизмографічного сигналу тривалістю 30 секунд.
Підготовлений PPG-сигнал	ppg_model	Фрагмент PPG-сигналу після заповнення пропусків, фільтрації, зменшення частоти дискретизації та нормалізації.

2.7. Розбиття даних на навчальну, валідаційну та тестову вибірки

Розбиття даних виконано за ідентифікатором клінічного кейсу. Це означає, що всі вікна одного кейсу потрапляють лише в одну вибірку: навчальну, валідаційну або тестову. Такий підхід є принципово важливим для коректної оцінки якості моделей.

Якщо виконати випадкове розбиття за окремими вікнами, виникне ризик витоку даних. Наприклад, фрагменти одного клінічного запису можуть опинитися і в навчальній, і в тестовій вибірці. Оскільки сусідні або близькі за часом вікна одного запису можуть бути дуже схожими, модель фактично тестуватиметься на даних, близьких до тих, які вона вже бачила під час навчання. Це призведе до завищення метрик і неправильної оцінки здатності моделі працювати на нових клінічних записах.

У роботі використано співвідношення 70% / 15% / 15% за клінічними кейсами. Початкове значення генератора випадкових чисел для розбиття становило 20260412. Кількісну характеристику отриманих навчальної, валідаційної та тестової вибірок наведено в таблиці 2.2.

Таблиця 2.2 – Характеристика навчальної, валідаційної та тестової вибірок

Вибірка	Кількість кейсів	Кількість вікон	Середня кількість вікон на кейс	Середня ЧСС, уд/хв	Середня частка валідних PPG-відліків
Навчальна	4310	1 481 565	343,75	72,04	0,9937
Валідаційна	923	326 792	354,05	71,82	0,9952
Тестова	924	320 648	347,02	72,59	0,9945

Дані таблиці 2.2 свідчать, що середня ЧСС і середня частка валідних PPG-відліків є близькими в усіх трьох вибірках, що підтримує коректність подальшого порівняння моделей.

Усього сформовано 2129005 вікон. Такий обсяг даних є достатнім для навчання моделей глибокого навчання, але одночасно вимагає обережної організації читання даних і обчислень.

2.8. Попередня обробка фотоплетизмографічного сигналу

Попередня обробка PPG-сигналу складається з кількох етапів. Її мета – привести сигнали з різних клінічних кейсів до спільного представлення, зменшити вплив шуму та зберегти інформацію, необхідну для оцінювання ЧСС.

Перший етап – заповнення пропущених значень. У реальних клінічних записах можуть траплятися пропуски або нечислові значення. Для коректної фільтрації такі пропуски потрібно обробити. У роботі використано лінійну інтерполяцію за наявними скінченними значеннями. Якщо пропуски розташовані на краях сигналу, вони заповнюються з урахуванням найближчих доступних значень.

Другий етап – смугова фільтрація. Для PPG-сигналу використано фільтр Баттерворта третього порядку, реалізований засобами пакета SciPy, з нижньою межею 0,5 Гц і верхньою межею 8,0 Гц [16]. Нижня межа потрібна для зменшення повільного дрейфу базової лінії. Верхня межа зменшує високочастотний шум. Обраний діапазон зберігає частотні компоненти, релевантні для пульсової хвилі та оцінювання ЧСС.

Третій етап – зменшення частоти дискретизації з 500 Гц до 125 Гц із використанням поліфазної ресемплінг-процедури [16]. Це зменшує розмір одного вікна з 15000 до 3750 відліків. Такий розмір є значно зручнішим для нейронних моделей і зберігає достатню часову роздільність для оцінювання серцевого ритму.

Четвертий етап – стійка нормалізація кожного вікна. Для цього використовується медіана та медіанне абсолютне відхилення. Нормалізований сигнал визначається як відхилення від медіани, поділене на стійку оцінку масштабу. Якщо масштаб є невалідним або занадто малим, використовується стандартне відхилення. Якщо й воно невалідне, повертається нульовий сигнал.

Стійка нормалізація потрібна тому, що абсолютна амплітуда PPG може сильно відрізнятися між пацієнтами, датчиками та періодами запису. Для оцінювання ЧСС важливішими є ритм і форма сигналу, ніж абсолютний масштаб. Нормалізація допомагає моделі зосередитися саме на цих властивостях.

Послідовність попередньої обробки PPG-сигналу наведено на рисунку 2.2. Рисунок 2.2 ілюструє, що попередня обробка складається з послідовних етапів заповнення пропусків, смугової фільтрації, зменшення частоти дискретизації та стійкої нормалізації.



Рисунок 2.2 – Послідовність попередньої обробки PPG-сигналу

2.9. Розрахунок показників якості PPG-вікон

Якість PPG-сигналу може суттєво впливати на точність оцінювання частоти серцевих скорочень. Тому для кожного 30-секундного вікна було розраховано набір показників, які характеризують повноту сигналу, його амплітудну варіативність, змінність після попередньої обробки та можливу наявність обмеження амплітуди.

До таких показників належать:

1. Частка валідних відліків PPG-сигналу. Цей показник визначає, яка частина значень у сирому PPG-вікні є скінченною та придатною для подальшої обробки. Низьке значення показника свідчить про те, що

значна частина сигналу була відсутня або невалідна. У фінальному оцінюванні використовувалися лише ті тестові вікна, для яких частка валідних відліків становила не менше 0,95.

2. Міжквартильний розмах амплітуди сирого PPG-сигналу. Цей показник характеризує типову варіативність амплітуди сигналу та є стійкішим до поодиноких викидів, ніж звичайний розмах між мінімальним і максимальним значеннями. Дуже низьке значення показника може свідчити про слабкий, майже плоский або недостатньо інформативний сигнал.
3. Стандартне відхилення підготовленого PPG-сигналу. Цей показник характеризує варіативність сигналу після заповнення пропусків, фільтрації, зменшення частоти дискретизації та нормалізації. Низьке значення показника може вказувати на недостатньо виражену пульсову структуру у відповідному вікні.
4. Частка потенційно обмежених за амплітудою значень. Цей показник оцінює, яка частина відліків сигналу знаходиться поблизу крайніх рівнів амплітуди. Підвищене значення може свідчити про можливе насичення, обрізання або інше апаратне обмеження сигналу.

Ці показники використовуються в роботі трьома способами. По-перше, вони допомагають фільтрувати вікна недостатньої якості. По-друге, вони входять до набору ознак для XGBoost. По-третє, вони використовуються для аналізу того, як якість сигналу впливає на похибку моделей.

2.10. Статистичний метод оцінювання ЧСС на основі піків PPG

Статистичний метод є базовим інтерпретованим підходом, з яким порівнюються навчальні моделі. Він не має навчальних параметрів і не потребує етапу навчання. Його логіка ґрунтується на тому, що кожна пульсова хвиля PPG приблизно відповідає одному серцевому скороченню.

Для кожного підготовленого PPG-вікна виконується така послідовність дій. Спочатку сигнал центрується шляхом віднімання медіани. Потім оцінюється

мінімальна виразність піків. Вона визначається як максимум між часткою стандартного відхилення сигналу та фіксованим мінімальним значенням. Далі виконується пошук піків із урахуванням мінімальної відстані між ними.

Мінімальна відстань між піками задається з урахуванням максимально допустимої ЧСС. У роботі використано верхню межу 180 уд/хв. Це означає, що два справжні пульсові піки не повинні бути ближчими, ніж інтервал, який відповідає такій частоті.

Після знаходження піків обчислюються міжпікові інтервали. Далі вони фільтруються за фізіологічно правдоподібним діапазоном, що відповідає ЧСС від 40 до 180 уд/хв. Фінальна оцінка ЧСС обчислюється як:

$$HR = 60 / \text{median}(IBI), \quad (2.2)$$

де IBI – міжпіковий інтервал у секундах.

Якщо вікно містить менше двох піків або після фільтрації не залишається валідних міжпікових інтервалів, прогноз вважається невалідним. У фінальному тестовому порівнянні статистичний метод мав 23 нечислові прогнози, тому спільна тестова вибірка для всіх моделей становила 318268 вікон.

Перевага методу – простота й висока швидкодія. Недолік – залежність від правильності пошуку піків. У результатах ця залежність проявляється у вигляді низької медіанної похибки, але довгого хвоста великих помилок.

2.11. Модель XGBoost

Модель XGBoost використовується як сильний підхід машинного навчання на основі ручних ознак [10]. На відміну від нейронних моделей, вона не отримує повний часовий ряд як вхід. Кожне PPG-вікно попередньо перетворюється на вектор із 133 ознак.

До набору ознак входять:

- 4 показники якості сигналу;
- 10 глобальних статистичних ознак;

- 7 процентилів;
- 48 локальних середніх значень;
- 48 локальних стандартних відхилень;
- 16 спектральних ознак.

Глобальні статистичні ознаки описують загальну форму сигналу: середнє значення, стандартне відхилення, мінімум, максимум, середнє абсолютне значення, середньоквадратичне значення, середнє значення похідної, стандартне відхилення похідної, середню абсолютну похідну та частоту зміни знаку.

Локальні ознаки отримуються шляхом поділу вікна на сегменти. Для кожного сегмента обчислюються середнє значення та стандартне відхилення. Це дозволяє частково зберегти часову структуру сигналу, хоча й у стиснутому вигляді.

Спектральні ознаки описують частотний склад сигналу. Вони є корисними, оскільки ЧСС пов'язана з періодичністю PPG-сигналу.

Для навчання XGBoost використано регресійну постановку з квадратичною функцією втрат. Структуру ознак і основні параметри моделі XGBoost наведено в таблиці 2.3. XGBoost є важливою проміжною моделлю між простим статистичним методом і нейронними мережами. Якщо статистичний метод використовує лише піки, то XGBoost використовує ширший опис сигналу. Водночас він усе ще залежить від ручного представлення, тому може втрачати частину інформації про форму PPG-хвилі.

Повна автоматизована оптимізація гіперпараметрів XGBoost у роботі не проводилася. Використана конфігурація розглядалася як регуляризоване робоче налаштування, обране з урахуванням великого обсягу даних, необхідності обмеження складності моделі та контролю перенавчання за валідаційною вибіркою.

Таблиця 2.3 – Структура та основні параметри моделі XGBoost

Компонент	Конфігурація	Призначення
Вхідне представлення	133 ручні ознаки	Компактний числовий опис 30-секундного PPG-вікна.
Показники якості сигналу	4 ознаки	Характеризують повноту, амплітудну варіативність, змінність після обробки та можливе обмеження сигналу.
Глобальні статистичні ознаки	10 ознак	Описують загальні статистичні властивості PPG-сигналу та його похідної.
Процентильні ознаки	7 ознак	Відображають розподіл амплітуди PPG-сигналу в межах одного вікна.
Локальні часові ознаки	96 ознак	Містять середні значення та стандартні відхилення в окремих часових сегментах вікна.
Спектральні ознаки	16 ознак	Описують частотну структуру PPG-сигналу.
Цільова функція	Середньоквадратична похибка	Використовується для навчання моделі в регресійній постановці.
Максимальна глибина дерева	4	Обмежує складність окремих дерев і зменшує ризик перенавчання.
Кількість дерев	700	Визначає максимальну кількість послідовно побудованих дерев у ансамблі.
Швидкість навчання	0,03	Контролює внесок кожного нового дерева в загальний прогноз моделі.
Частка рядків	0,85	Використовується для стохастичного навчання та додаткової регуляризації.

Закінчення таблиці 2.3

Компонент	Конфігурація	Призначення
Частка ознак	0,85	Обмежує частину ознак, доступних для побудови окремих дерев.
Регуляризація	L1 = 0,05; L2 = 3,0	Зменшує складність моделі та ризик перенавчання.
Рання зупинка	40 раундів без покращення	Завершує навчання, якщо якість на валідаційній вибірці не покращується.

2.12. Архітектура нейронної моделі Conv+LSTM

Модель Conv+LSTM є основною нейронною моделлю роботи. Її було реалізовано засобами PyTorch і побудовано на поєднанні одновимірних згорткових шарів та двонапрявленого LSTM-блоку [11, 17]. Вона працює з підготовленим PPG-вікном довжиною 3750 відліків і формує прогноз ЧСС.

Архітектура складається з трьох основних частин:

- згортковий блок;
- двонапрявлений LSTM-блок;
- регресійна голова.

Згортковий блок містить чотири одновимірні згорткові шари. Кожен шар поступово зменшує довжину послідовності та збільшує кількість каналів ознак. Після згортки використовуються нормалізація, нелінійна активація та проріджування. Такий блок виділяє локальні патерни PPG-сигналу: підйоми, спади, піки, локальні деформації та короткочасні зміни.

Після згорткового блоку послідовність ознак передається до двонапрявленого LSTM. Його задача – узагальнити часову структуру всього вікна. Оскільки вікно аналізується повністю, модель може використовувати інформацію як з початку, так і з кінця фрагмента. Саме тому використовується двонапрявлений варіант.

Після LSTM виконується усереднення за часовою віссю. Отриманий вектор передається до регресійної голови, яка складається з нормалізації, повнозв’язного шару, нелінійної активації, проріджування та фінального шару з одним виходом.

Модель має 57121 навчальний параметр і розмір приблизно 237 КБ. Така компактність є важливою перевагою, оскільки модель може бути потенційно використана в системах із обмеженими ресурсами. Архітектуру та основні параметри моделі Conv+LSTM наведено в таблиці 2.4.

Таблиця 2.4 – Архітектура Conv+LSTM

Компонент	Конфігурація	Призначення
Вхідний сигнал	30-секундне PPG-вікно, 3750 відліків, 1 канал	Підготовлений PPG-сигнал після заповнення пропусків, фільтрації, зменшення частоти дискретизації та стійкої нормалізації.
Згортковий блок 1	1D-згортка 1→16, ядро 9, крок 2; BatchNorm; SiLU	Початкове виділення локальних ознак форми PPG-хвилі.
Згортковий блок 2	1D-згортка 16→24, ядро 7, крок 2; BatchNorm; SiLU	Подальше стискання часової послідовності та розширення простору ознак.
Згортковий блок 3	1D-згортка 24→32, ядро 5, крок 2; BatchNorm; SiLU	Формування більш узагальнених локальних характеристик сигналу.
Згортковий блок 4	1D-згортка 32→48, ядро 5, крок 2; BatchNorm; SiLU	Отримання компактного послідовного представлення PPG-вікна.
Проріджування	Dropout = 0,15	Регуляризація моделі та зменшення ризику перенавчання.

Продовження таблиці 2.4

Компонент	Конфігурація	Призначення
Рекурентний блок	Двонаправлений LSTM, розмір прихованого стану = 48, 1 шар	Узагальнення ритмічної та часової структури сигналу в межах 30-секундного вікна.
Агрегація	Усереднення за часовою віссю	Перетворення послідовності ознак в один вектор представлення.
Регресійна голова	LayerNorm; повнозв'язний шар 96→48; SiLU; Dropout; повнозв'язний шар 48→1	Формування фінального числового прогнозу ЧСС.
Вихід моделі	Одне числове значення	Прогнозована частота серцевих скорочень, уд/хв.
Кількість параметрів	57121	Загальна кількість навчальних параметрів моделі.
Розмір моделі	приблизно 237 КБ	Орієнтовний розмір збереженої моделі.

Навчання виконувалося з нормалізацією цільової змінної. Модель прогнозувала нормалізоване значення ЧСС, а після прогнозування результат перетворювався назад в уд/хв. Це допомагає стабілізувати навчання та покращити числову поведінку моделі.

2.13. Архітектура нейронної моделі Conv+Transformer

Модель Conv+Transformer є другою нейронною архітектурою, використаною для порівняння. Її було реалізовано засобами PyTorch і побудовано на основі згорткового перетворення сигналу та кодувальника Transformer [12, 17].

Перший етап моделі – згорткове перетворення сигналу. Воно зменшує довжину послідовності та перетворює сирі відліки в послідовність векторних ознак. Далі до цих ознак додається позиційне представлення, оскільки механізм уваги сам по собі не містить інформації про порядок елементів.

Після цього послідовність обробляється кодувальником Transformer. У роботі використано двошаровий кодувальник із чотирма головами уваги. Механізм уваги дозволяє моделі враховувати взаємозв'язки між різними ділянками сигналу. Це може бути корисним, якщо важливі патерни розташовані на відстані один від одного в межах вікна.

Після кодувальника виконується усереднення за часовою віссю, а потім регресійна голова формує прогноз ЧСС. Архітектуру та основні параметри моделі Conv+Transformer наведено в таблиці 2.5.

Таблиця 2.5 – Архітектура Conv+Transformer

Компонент	Конфігурація	Призначення
Вхідний сигнал	30-секундне PPG-вікно, 3750 відліків, 1 канал	Підготовлений PPG-сигнал після попередньої обробки.
Згортковий блок 1	1D-згортка 1→16, ядро 9, крок 2; BatchNorm; GELU	Початкове виділення локальних ознак PPG-сигналу.
Згортковий блок 2	1D-згортка 16→32, ядро 7, крок 2; BatchNorm; GELU	Стискання часової послідовності та збільшення кількості ознак.
Згортковий блок 3	1D-згортка 32→48, ядро 5, крок 2; BatchNorm; GELU	Формування послідовності векторних ознак для Transformer-кодувальника.
Додаткове стискання	Середнє об'єднання з ядром 3 і кроком 3	Зменшення довжини послідовності перед застосуванням механізму уваги.
Позиційне представлення	Навчуване позиційне кодування	Додавання інформації про порядок елементів у послідовності.
Transformer Encoder	2 шари, розмір представлення 48, 4 голови уваги, feed-forward розмір 96	Моделювання взаємозв'язків між різними ділянками PPG-вікна.
Проріджування	Dropout = 0,15	Регуляризація моделі та зменшення ризику перенавчання.

Продовження таблиці 2.5

Компонент	Конфігурація	Призначення
Агрегація	Усереднення за часовою віссю	Перетворення послідовності ознак в один вектор представлення.
Регресійна голова	LayerNorm; повнозв'язний шар 48→48; GELU; Dropout; повнозв'язний шар 48→1	Формування фінального числового прогнозу ЧСС.
Вихід моделі	Одне числове значення	Прогнозована частота серцевих скорочень, уд/хв.
Кількість параметрів	59745	Загальна кількість навчальних параметрів моделі.
Розмір моделі	приблизно 252,1 КБ	Орієнтовний розмір збереженої моделі.

Conv+Transformer є корисною для порівняння з Conv+LSTM, оскільки ці моделі використовують різні способи узагальнення часової інформації. У результатах роботи Conv+Transformer показала високу якість, але поступилася Conv+LSTM за середньою абсолютною похибкою.

2.14. Організація навчання моделей

Навчання моделей виконувалося на навчальній вибірці, а контроль якості – на валідаційній. Тестова вибірка використовувалася лише після завершення навчання для фінального оцінювання.

Для нейронних моделей використовувалося потокове читання даних. Це необхідно через великий обсяг сформованого набору даних. Замість завантаження всіх вікон в оперативну пам'ять дані читалися пакетами. Для кожного пакета виконувалася перевірка валідності сигналу й цільового значення.

Під час навчання нейронних моделей використовувалися:

- розмір пакета 128;
- оптимізація з регуляризацією ваг;
- рання зупинка за валідаційною якістю;

- обмеження норми градієнта;
- нормалізація цільової змінної.

Тренування й тестування моделей виконувалося на обчислювальному середовищі з процесором Apple Silicon M5 Pro.

Для XGBoost навчання виконувалося на векторах ручних ознак. Якість контролювалася на валідаційній вибірці, а для запобігання перенавчанню використовувалася рання зупинка.

Статистичний метод на основі пошуку піків PPG-сигналу не потребував навчання. Для нього одразу формувалися прогнози за заданим алгоритмом пошуку піків.

2.15. Метрики оцінювання якості прогнозування

Для оцінювання моделей використано комплекс метрик, оскільки жодна окрема метрика не описує якість повністю.

Середня абсолютна похибка визначається як середнє значення модуля різниці між прогнозом і референсом. Середня абсолютна похибка визначається за формулою (2.3).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \quad (2.3)$$

Ця метрика є основною, оскільки вимірюється в уд/хв і легко інтерпретується.

Корінь середньоквадратичної похибки визначається як:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (2.4)$$

RMSE сильніше штрафує великі помилки, тому корисна для виявлення моделей із довгим хвостом похибок.

Систематичний зсув визначається як:

$$Bias = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i), \quad (2.5)$$

Якщо зсув додатний, модель у середньому завищує ЧСС. Якщо від'ємний – занижує.

Коефіцієнт детермінації визначає частку варіативності референсної ЧСС, яку пояснює модель:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (2.6)$$

Також використовуються медіанна абсолютна похибка, 90-й і 95-й процентилі абсолютної похибки та частки прогнозів із похибкою не більше 3, 5, 10 і 15 уд/хв.

Порогові метрики мають практичну інтерпретацію. Наприклад, частка прогнозів у межах 5 уд/хв показує, яка частина результатів є близькою до референсу з погляду абсолютної різниці.

2.16. Оцінювання невизначеності та аналіз на рівні клінічних кейсів

Оскільки вікна одного клінічного запису корельовані між собою, для оцінювання невизначеності використано бутстреп за клінічними кейсами. У цьому підході повторне вибирання виконується не за окремими вікнами, а за ідентифікатором клінічного кейсу.

Процедура така:

- формується список тестових клінічних кейсів;
- на кожній ітерації випадково вибираються кейси з повторенням;

- для вибраних кейсів беруться всі відповідні вікна;
- на цих вікнах обчислюються метрики;
- після 1000 ітерацій визначаються 2,5-й і 97,5-й процентилі метрик.

Такий підхід дає більш реалістичні довірчі інтервали, ніж бутстреп за окремими вікнами. Він відповідає практичному питанню: як модель поводитиметься на нових клінічних записах.

2.17. Висновки до розділу 2

У другому розділі було сформульовано постановку задачі оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом як задачі регресії за 30-секундним часовим рядом. Визначено об'єкт, предмет, мету та завдання кваліфікаційної роботи, а також описано основні обмеження дослідження, пов'язані з використанням одного вхідного PPG-каналу, неідеальністю референсного каналу Solar8000/HR та відсутністю зовнішньої клінічної валідації.

Було обґрунтовано вибір відкритого набору клінічних даних VitalDB як джерела фізіологічних сигналів. Як вхідний сигнал використано канал SNUADC/PLETH, а як референсне значення частоти серцевих скорочень — канал Solar8000/HR. Такий вибір дозволив сформулювати задачу контрольованого навчання, у якій кожному 30-секундному фрагменту PPG-сигналу відповідає числове значення ЧСС.

Описано процес формування набору даних із неперекривних 30-секундних вікон. Для кожного вікна зберігалися ідентифікатор клінічного кейсу, часові межі, референсна ЧСС, показники якості сигналу, сирий PPG-сигнал і підготовлений PPG-сигнал. Розбиття даних на навчальну, валідаційну та тестову вибірки виконано за ідентифікатором клінічного кейсу у співвідношенні 70% / 15% / 15%, що дозволило уникнути витоку даних між вибірками.

Було розроблено послідовність попередньої обробки PPG-сигналу, яка включає заповнення пропущених значень, смугову фільтрацію, зменшення частоти дискретизації з 500 Гц до 125 Гц і стійку нормалізацію. Додатково для

кожного вікна обчислено показники якості сигналу: частку валідних відліків, міжквартильний розмах амплітуди, стандартне відхилення підготовленого сигналу та частку потенційно обмежених за амплітудою значень.

У розділі також описано чотири підходи до оцінювання ЧСС: класичний статистичний метод на основі піків PPG-сигналу, модель XGBoost на основі ручних ознак, нейронну модель Conv+LSTM та нейронну модель Conv+Transformer. Для кожного підходу визначено вхідне представлення даних, основну логіку роботи та роль у порівняльному експерименті.

Для оцінювання якості моделей обрано комплекс метрик, зокрема MAE, RMSE, R^2 , систематичний зсув, медіанну абсолютну похибку, процентилі абсолютної похибки та частку прогнозів у межах заданих порогів. Окремо описано оцінювання невизначеності за допомогою бутстрепа на рівні клінічних кейсів, що є важливим для коректної інтерпретації результатів на медичних часових рядах.

Отже, у другому розділі було розроблено повну методику створення програмного засобу для автоматизованого оцінювання ЧСС за PPG-сигналом: від вибору джерела даних і формування вікон до попередньої обробки, побудови моделей, організації навчання та вибору метрик експериментального оцінювання.

РОЗДІЛ 3. ЕКСПЕРИМЕНТАЛЬНІ РЕЗУЛЬТАТИ ТА АНАЛІЗ ЯКОСТІ РОЗРОБЛЕНИХ МОДЕЛЕЙ

3.1. Характеристика сформованого набору даних

У результаті виконання роботи було сформовано набір даних для задачі оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом. Джерелом даних були клінічні записи VitalDB. Для кожного клінічного кейсу використовувався PPG-канал SNUADC/PLETH і референсний канал частоти серцевих скорочень Solar8000/HR.

Фінальний набір даних містить 6157 клінічних кейсів. Кожен кейс було розбито на 30-секундні неперекривні вікна. Загальна кількість сформованих вікон становить 2129005. Такий обсяг дозволив навчати як моделі на основі ручних ознак, так і нейронні моделі, що працюють із повним часовим рядом.

Розбиття на навчальну, валідаційну й тестову вибірки виконано за ідентифікатором клінічного кейсу. Це означає, що всі вікна одного запису потрапляють лише в одну частину даних. Такий підхід був обраний для запобігання витоку інформації між навчанням і тестуванням.

Сформований розподіл даних наведено в таблиці 3.1.

Таблиця 3.1 – Підсумкова характеристика сформованого набору даних

Вибірка	Кількість кейсів	Кількість вікон	Середня кількість вікон на кейс	Середня ЧСС, уд/хв	Середня частка валідних PPG-відліків
Навчальна	4310	1481565	343,75	72,04	0,9937
Валідаційна	923	326792	354,05	71,82	0,9952
Тестова	924	320648	347,02	72,59	0,9945

З таблиці видно, що середня частка валідних PPG-відліків є високою в усіх вибірках і перевищує 0,99. Це означає, що більшість сформованих вікон містять достатньо повний сигнал. Водночас навіть невелика частка невалідних або низькоякісних вікон може впливати на результати, тому для фінального оцінювання застосовано додатковий фільтр якості.

У тестовій вибірці початково було 320648 вікон. Після фільтрації за критерієм частки валідних відліків PPG-сигналу не менше 0,95 залишилося 318291 вікно. Для парного порівняння всіх моделей використовувалися лише ті вікна, для яких усі методи сформували скінченні прогнози. Оскільки базовий статистичний метод на основі піків PPG-сигналу не сформував числовий прогноз для 23 вікон, фінальна спільна тестова вибірка становила 318268 вікон.

Аудит тестової вибірки після фільтрації якості наведено в таблиці 3.2.

Таблиця 3.2 – Аудит тестової вибірки після фільтрації якості

Показник	Значення
Усі тестові вікна	320648
Тестові вікна після фільтра якості	318291
Відкинуті низьковалідні вікна	2357
Кількість тестових кейсів після фільтрації	924
Спільні тестові вікна для всіх моделей	318268

З таблиці 3.2 видно, що після фільтрації за якістю та вилучення вікон із нечисловими прогнозами статистичного методу для парного порівняння залишилося 318268 тестових вікон. Це забезпечує коректність порівняння, оскільки всі моделі оцінювалися на однакових прикладах.

3.2. Розподіл референсної частоти серцевих скорочень

Для інтерпретації результатів важливо проаналізувати розподіл референсної частоти серцевих скорочень у тестовій вибірці. На якісних тестових вікнах він мав такі характеристики:

Таблиця 3.3 – Розподіл референсної ЧСС у тестовій вибірці

Показник	Значення, уд/хв
Середнє	72,59
Стандартне відхилення	15,26
Мінімум	0
1-й процентиль	46,33
5-й процентиль	52,20
Медіана	70,53
75-й процентиль	81,07
95-й процентиль	100,00
99-й процентиль	117,93
Максимум	300

Основна маса даних знаходиться приблизно в діапазоні 50–100 уд/хв. Медіана становить 70,53 уд/хв, а 95-й процентиль – 100 уд/хв. Це означає, що більшість прикладів відповідає центральному фізіологічному діапазону.

Водночас у даних є екстремальні значення 0 і 300 уд/хв. Такі значення можуть бути результатом артефактів референсного каналу або особливостей моніторної системи. Їх наявність важлива для оцінювання моделей, оскільки вони можуть суттєво збільшувати RMSE і впливати на аналіз похибок у крайніх діапазонах.

Нерівномірність розподілу ЧСС також означає, що загальні метрики значною мірою відображають якість у центральному діапазоні, де даних найбільше. Для чесної оцінки необхідно окремо аналізувати поведінку моделей у низьких, високих і екстремальних діапазонах. Такий аналіз наведено в підрозділі 3.8.

Візуально розподіл референсної ЧСС у тестовій вибірці наведено на рисунку 3.1.

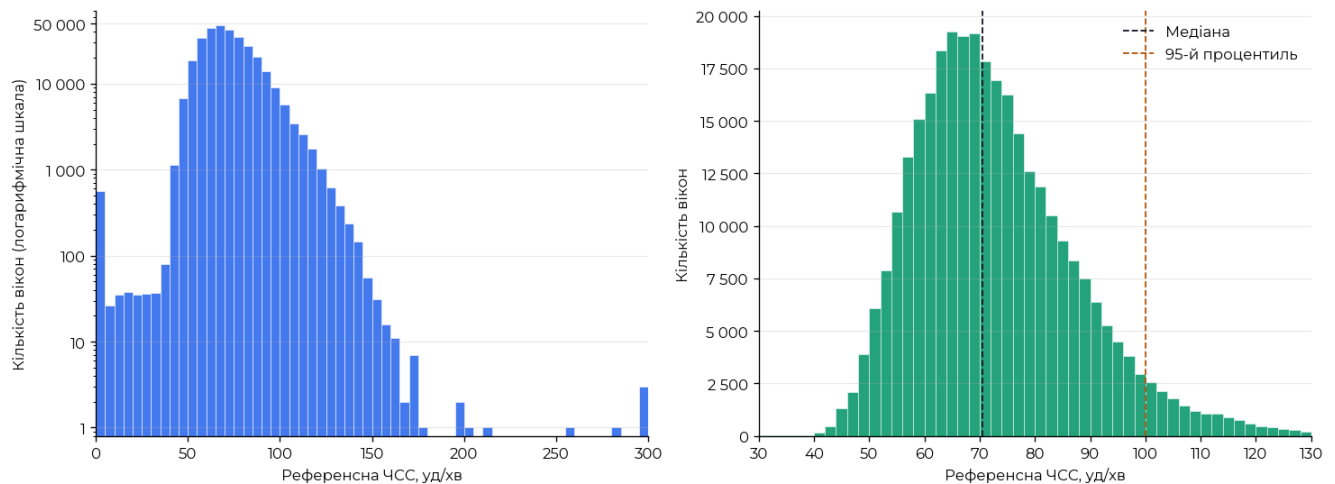


Рисунок 3.1 – Гістограма розподілу референсної ЧСС у тестовій вибірці

Як видно з рисунка 3.1, більшість вікон зосереджена в центральному діапазоні ЧСС, тоді як крайні значення трапляються значно рідше.

3.3. Основне порівняння моделей на спільній тестовій вибірці

Основне порівняння виконано на 318268 спільних тестових вікнах. Було порівняно чотири підходи. Результати наведено в таблицях 3.4 і 3.5.

Таблиця 3.4 – Основні метрики якості моделей на тестовій вибірці

Модель	MAE, уд/хв	95% ДІ MAE	RMSE, уд/хв	95% ДІ RMSE	Зсув, уд/хв	R ²
Conv+LSTM	1,803	[1,657; 1,987]	6,075	[5,53; 6,72]	-0,448	0,84
Conv+Transformer	3,151	[3,00; 3,35]	6,653	[6,08; 7,41]	-1,985	0,81
XGBoost	3,782	[3,56; 4,06]	7,588	[6,95; 8,45]	-0,038	0,75
Статистичний метод PPG	5,183	[4,72; 5,68]	15,164	[14,11; 16,17]	+3,841	0,01

Таблиця 3.5 – Розподільчі характеристики абсолютної похибки моделей

Модель	Медіанна AE, уд/хв	P95 AE, уд/хв	AE ≤ 5 уд/хв	AE ≤ 10 уд/хв
Conv+LSTM	0,493	6,579	93,64%	96,59%
Conv+Transformer	1,972	8,621	89,50%	95,80%
XGBoost	1,864	13,069	79,54%	92,31%
Статистичний метод PPG	0,770	38,940	86,80%	90,15%

Примітка. ДІ – довірчий інтервал; AE – абсолютна похибка; MAE, RMSE, медіанна AE, P95 AE і зсув наведено в уд/хв.

Найкращі результати показала модель Conv+LSTM. Вона має найнижчу середню абсолютну похибку, найнижчу медіанну абсолютну похибку, найнижчий 95-й перцентиль абсолютної похибки, найвище значення коефіцієнта детермінації та найбільшу частку прогнозів у межах 5 і 10 уд/хв.

Conv+Transformer посіла друге місце за MAE. Вона має кращу якість за XGBoost за середньою абсолютною похибкою і пороговими метриками, але має помітніший від'ємний зсув. Це означає, що модель у середньому занижує ЧСС приблизно на 1,985 уд/хв.

XGBoost має майже нульовий середній зсув, що є позитивною властивістю. Однак за MAE, P95 AE і часткою прогнозів у межах 5 уд/хв вона поступається нейронним моделям. Це свідчить про те, що ручні ознаки є інформативними, але не повністю замінюють роботу з повною часовою формою PPG-сигналу.

Статистичний метод PPG має характерну подвійну поведінку. Його медіанна абсолютна похибка становить лише 0,770 уд/хв, що свідчить про високу точність на багатьох простих вікнах. Однак RMSE дорівнює 15,164 уд/хв, а P95 AE – 38,940 уд/хв. Це означає, що модель має довгий хвіст великих похибок. Саме такі великі помилки роблять статистичний підхід менш надійним для загального використання.

Крім основних метрик, було розраховано коефіцієнти кореляції між прогнозованою та референсною ЧСС. Коефіцієнти кореляції між прогнозованою та референсною ЧСС наведено в таблиці 3.6.

Таблиця 3.6 – Кореляція між прогнозованою та референсною ЧСС

Модель	Кореляція Пірсона	Кореляція Спірмена
Статистичний метод PPG	0,642	0,808
XGBoost	0,868	0,913
Conv+LSTM	0,918	0,952
Conv+Transformer	0,911	0,946

Найвищі кореляції має Conv+LSTM. Це означає, що ця модель найкраще відтворює загальну зміну ЧСС між вікнами. Conv+Transformer має близькі, але дещо нижчі значення. XGBoost також демонструє сильну кореляцію, але

поступається нейронним моделям. Статистичний метод PPG має нижчу кореляцію, що пов'язано з великою кількістю викидів похибки.

3.4. Аналіз розподілу абсолютної похибки

Середні метрики дають загальне уявлення про якість моделей, але не показують повну форму розподілу похибок. Тому окремо було проаналізовано емпіричний розподіл абсолютної похибки.

Conv+LSTM найшвидше набирає частку малих похибок. Це означає, що для більшості тестових вікон її прогнози знаходяться дуже близько до референсної ЧСС. 93,64% прогнозів цієї моделі мають похибку не більше 5 уд/хв, а 96,59% – не більше 10 уд/хв.

Conv+Transformer також має добрий розподіл похибок: 89,50% вікон мають похибку не більше 5 уд/хв, а 95,80% – не більше 10 уд/хв. Однак порівняно з Conv+LSTM її крива абсолютної похибки зміщена в бік більших значень.

XGBoost має гіршу частку прогнозів у межах 5 уд/хв – 79,54%. Водночас 92,31% прогнозів мають похибку не більше 10 уд/хв, тобто модель залишається достатньо стабільною для багатьох вікон, але менш точною в зоні малих похибок.

Статистичний метод PPG має високу частку точних прогнозів на простих вікнах, але його крива має довгий хвіст. Це означає, що частина прогнозів має дуже велику похибку. Така поведінка є типовою для методів, які залежать від одного критичного етапу – правильного пошуку піків.

Емпіричну функцію розподілу абсолютної похибки та частки прогнозів у межах порогів 5 і 10 уд/хв наведено на рисунку 3.2. З практичного погляду важливо, що Conv+LSTM не лише має найкращу середню похибку, а й забезпечує найбільшу частку прогнозів у межах малих порогів. Це означає, що перевага моделі не є результатом лише окремих дуже точних прогнозів, а проявляється в усьому розподілі похибки.

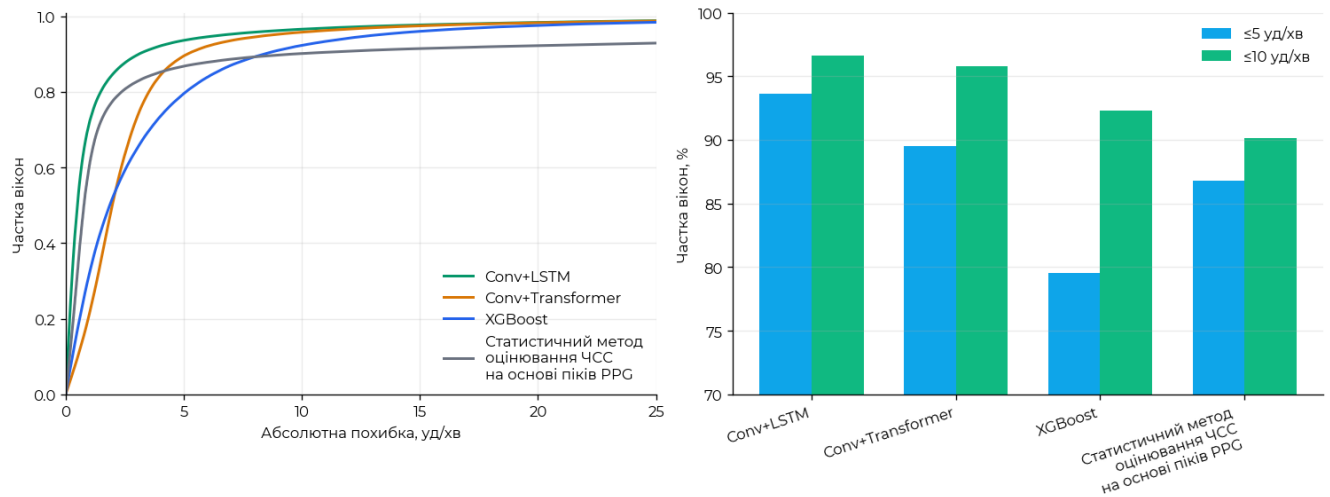


Рисунок 3.2 – Емпірична функція розподілу абсолютної похибки та частки прогнозів у межах 5 і 10 уд/хв

3.5. Аналіз результатів на рівні клінічних кейсів

Оскільки вікна одного клінічного кейсу корельовані, окремо було виконано аналіз на рівні кейсів. Для кожного тестового кейсу було обчислено MAE кожної моделі, після чого розраховано підсумкові статистики. Згідно з таблицею 3.7, Conv+LSTM має найнижчу медіану MAE на рівні клінічних кейсів – 1,142 уд/хв. Це означає, що для типового клінічного кейсу модель помиляється приблизно на 1,14 уд/хв у середньому за вікнами цього кейсу. 95-й перцентиль MAE на рівні клінічних кейсів становить 5,885 уд/хв, що також є найкращим результатом серед усіх моделей.

Conv+Transformer займає друге місце за показниками на рівні клінічних кейсів. XGBoost поступається обом нейронним моделям, але має кращий хвіст складних кейсів, ніж статистичний метод PPG. Статистичний підхід має відносно непогану медіану, але дуже високі середнє значення та P95 MAE. Це ще раз підтверджує, що його головна проблема – нестабільність на складних записах.

Таблиця 3.7 – Порівняння моделей на рівні клінічних кейсів

Модель	Медіана MAE на рівні кейсів	Середня MAE на рівні кейсів	P90 MAE на рівні кейсів	P95 MAE на рівні кейсів
Conv+LSTM	1,142	1,937	3,869	5,885
Conv+Transformer	2,556	3,234	5,016	7,227
XGBoost	3,057	3,888	6,584	9,251
Статистичний метод PPG	2,358	5,389	13,842	21,433

Парні статистичні тести за MAE на рівні клінічних кейсів показали, що Conv+LSTM має статистично переконливу перевагу над статистичним методом, XGBoost і Conv+Transformer. Це важливо, оскільки перевага моделі підтверджується не лише на рівні окремих вікон, а й на рівні клінічних записів.

Розподіл MAE на рівні клінічних кейсів для досліджуваних моделей наведено на рисунку 3.3. Аналіз на рівні клінічних кейсів є одним із найважливіших результатів роботи. Він показує, що Conv+LSTM не просто добре працює на великій кількості окремих вікон, а стабільно демонструє найкращу якість на більшості клінічних кейсів. Це суттєво підсилює висновок про її практичну перевагу.

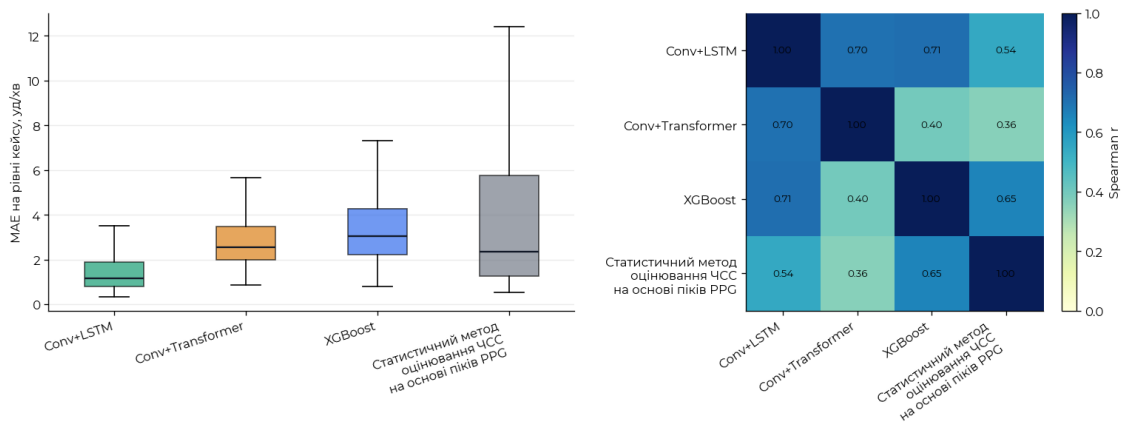


Рисунок 3.3 – Розподіл MAE моделей на рівні клінічних кейсів

3.6. Аналіз узгодженості прогнозів методом Бленда-Альтмана

Для оцінювання узгодженості прогнозованої та референсної ЧСС було використано аналіз Бленда-Альтмана. Він дозволяє оцінити середній зсув прогнозів і межі узгодження. Підсумкові числові характеристики аналізу Бленда-Альтмана наведено в таблиці 3.8.

Таблиця 3.8 – Підсумок аналізу Бленда-Альтмана

Модель	Зсув, уд/хв	Стандартне відхилення похибки	Межі узгодження, уд/хв
Статистичний метод PPG	+3,841	14,669	[-24,91; +32,59]
XGBoost	-0,038	7,588	[-14,91; +14,83]
Conv+LSTM	-0,448	6,059	[-12,32; +11,43]
Conv+Transformer	-1,985	6,350	[-14,43; +10,46]

З таблиці 3.8 видно, що Conv+LSTM має найвужчі межі узгодження та невеликий від’ємний зсув. Це означає, що її прогнози найбільш компактно зосереджені навколо референсного значення.

XGBoost має майже нульовий середній зсув, але ширші межі узгодження. Це означає, що модель у середньому не завищує і не занижує ЧСС, але її окремі похибки мають більший розкид.

Conv+Transformer має помітний від’ємний зсув, тобто схильність до заниження ЧСС. Статистичний метод PPG має додатний зсув і найширші межі узгодження, що пов’язано з великою кількістю викидів похибки. Графіки Бленда-Альтмана для чотирьох моделей наведено на рисунку 3.4.

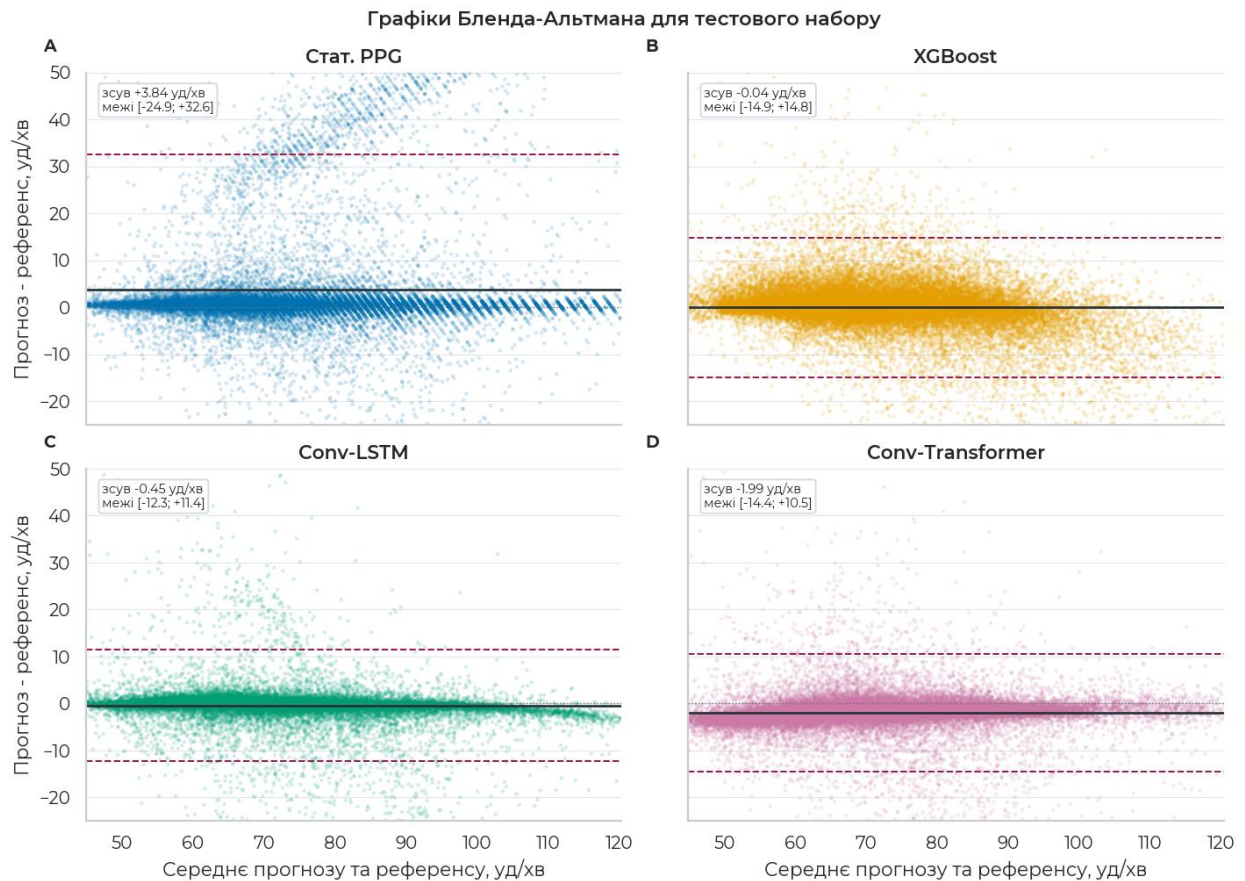


Рисунок 3.4 – Графіки Бленда-Альтмана для чотирьох моделей

Аналіз Бленда-Альтмана підтверджує результати основного порівняння. Conv+LSTM не лише має найнижчу середню абсолютну похибку, а й демонструє найкраще узгодження з референсною ЧСС. Це важливо, оскільки для практичних задач недостатньо мати низьку середню похибку; потрібно також, щоб різниця між прогнозом і референсом була стабільною та не мала великого розкиду.

3.7. Калібраційний аналіз прогнозованої та референсної ЧСС

Калібраційний аналіз виконувався за допомогою графіків залежності прогнозованої ЧСС від референсної. Ідеальна модель мала б формувати точки вздовж діагональної лінії, де прогноз дорівнює референсу.

Conv+LSTM формує найщільніше скупчення навколо діагоналі. Це узгоджується з її найкращими значеннями MAE, R^2 і порогових метрик. Модель добре відтворює центральний діапазон ЧСС, де знаходиться основна маса даних. Conv+Transformer також демонструє виражену діагональну структуру, але має

помітний зсув у бік нижчих прогнозів. Це відповідає від'ємному середньому зсуву. XGBoost має ширше скупчення точок. У центральному діапазоні модель працює стабільно, але на високих значеннях ЧСС помітна тенденція до заниження. Статистичний метод PPG має багато точних попадань, але також багато розсіяних точок далеко від діагоналі. Це відображає його нестабільність у складних вікнах, де пошук піків дає неправильну оцінку ритму.

Калібраційні графіки прогнозованої та референсної ЧСС наведено на рисунку 3.5. Калібраційний аналіз важливий тим, що він дозволяє виявити не лише середню помилку, а й характер поведінки моделі в різних областях значень. У цій роботі видно, що головною проблемою навчальних моделей є заниження високих значень ЧСС. Саме тому додатково проведено аналіз похибки за діапазонами референсної ЧСС.

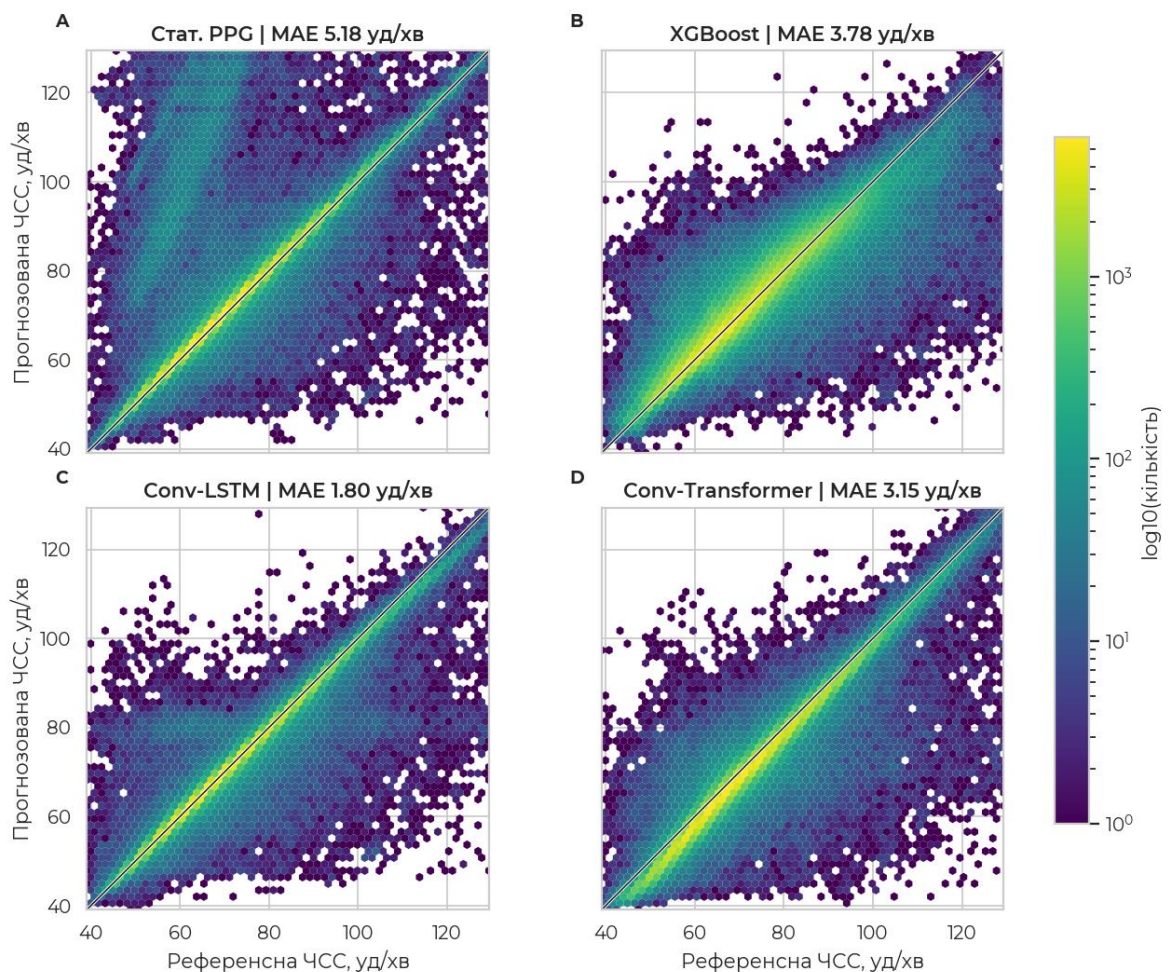


Рисунок 3.5 – Калібраційні графіки прогнозованої та референсної ЧСС

3.8. Аналіз похибки за діапазонами ЧСС

Для аналізу залежності якості моделей від референсної ЧСС тестові вікна було поділено на діапазони. Результати наведено в таблиці 3.9. Найкращі результати Conv+LSTM демонструє у центральному діапазоні 50–90 уд/хв, де зосереджена більшість даних. У цих діапазонах MAE моделі часто знаходиться близько 1–2 уд/хв. Це свідчить про високу якість моделі для типових значень ЧСС.

Таблиця 3.9 – MAE моделей за діапазонами референсної ЧСС

Діапазон ЧСС	Кількість вікон	Статистичний метод PPG	XGBoost	Conv +LSTM	Conv +Transformer
<40	849	77,50	63,69	64,59	60,99
40–50	7888	7,08	4,49	2,73	4,04
50–60	52960	6,37	2,85	1,22	3,34
60–70	92141	5,20	2,63	1,04	2,40
70–80	78046	4,78	2,82	1,13	2,36
80–90	47496	4,20	3,33	1,62	2,75
90–100	22875	2,45	5,02	2,40	3,43
100–120	13450	3,73	10,96	5,05	5,64
120–160	2533	11,52	30,48	18,09	19,39
≥160	30	92,19	106,93	106,12	108,48

У діапазоні 100–120 уд/хв похибка зростає. Для Conv+LSTM вона становить 5,05 уд/хв, для Conv+Transformer – 5,64 уд/хв, для XGBoost – 10,96

уд/хв. Це свідчить, що на підвищених значеннях ЧСС моделі починають працювати гірше, особливо XGBoost.

У діапазоні 120–160 уд/хв усі навчальні моделі мають суттєве погіршення. Conv+LSTM залишається кращою серед навчальних моделей, але її MAE зростає до 18,09 уд/хв. Для XGBoost похибка становить 30,48 уд/хв. Це свідчить про тенденцію моделей до заниження високих значень ЧСС.

Діапазон ≥ 160 уд/хв містить лише 30 вікон, тому висновки для нього є нестійкими. Усі моделі мають дуже велику похибку. Частково це може бути пов'язано з малою кількістю таких прикладів, а частково – з артефактами референсного каналу.

Залежність MAE від діапазону референсної ЧСС наведено на рисунку 3.6.

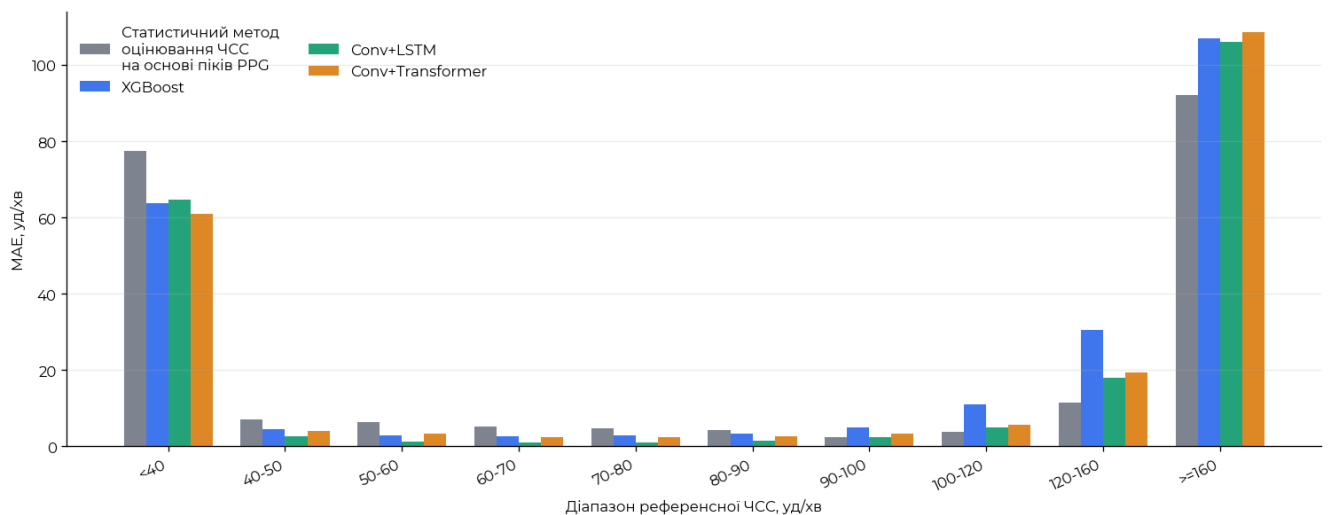


Рисунок 3.6 – Залежність MAE від діапазону референсної ЧСС

Як видно з рисунка 3.6, найнижча похибка спостерігається в центральному діапазоні ЧСС, тоді як у рідкісних крайніх діапазонах похибка різко зростає. Це є одним із важливих напрямків подальшого вдосконалення.

3.9. Аналіз впливу якості PPG-сигналу на похибку

Якість PPG-сигналу є одним із ключових чинників, що впливають на точність оцінювання ЧСС. У роботі було проаналізовано похибку моделей за

квартилями двох показників: стандартного відхилення підготовленого PPG і міжквартильного розмаху амплітуди сирого PPG.

За квартилями стандартного відхилення PPG результати були такими:

Таблиця 3.10 – MAE моделей за квартилями стандартного відхилення підготовленого PPG-сигналу

Квартиль	Статистичний метод PPG	XGBoost	Conv+LSTM	Conv+Transformer
Q1, найнижчий	7,10	5,07	2,55	3,74
Q2	5,39	3,87	1,70	3,21
Q3	5,04	2,67	1,23	2,74
Q4, найвищий	3,19	3,52	1,74	2,91

Таблиця 3.11 – MAE моделей за квартилями міжквартильного розмаху амплітуди сирого PPG-сигналу

Квартиль	Статистичний метод PPG	XGBoost	Conv+LSTM	Conv+Transformer
Q1, найнижчий	7,70	4,83	2,16	3,32
Q2	6,90	3,67	1,53	2,89
Q3	3,41	3,29	1,71	3,12
Q4, найвищий	2,73	3,35	1,81	3,27

Низька якість PPG погіршує результати всіх моделей, але ступінь погіршення різний. Статистичний метод PPG є найбільш чутливим до якості сигналу. Це очікувано, оскільки пошук піків прямо залежить від чіткості пульсової хвилі.

Conv+LSTM демонструє найстабільнішу поведінку. Навіть у найгірших квартилях якості її MAE залишається на рівні приблизно 2,16–2,55 уд/хв. Це

свідчить про те, що нейронна модель краще переносить деформації або слабку амплітуду сигналу.

XGBoost погіршується на вікнах із низьким стандартним відхиленням підготовленого сигналу, оскільки ручні ознаки стають менш інформативними. Conv+Transformer також має погіршення на найгірших квантилях, але воно менш різке, ніж у статистичного методу.

Візуальне порівняння похибки моделей за квантилями показників якості PPG-сигналу наведено на рисунку 3.7.

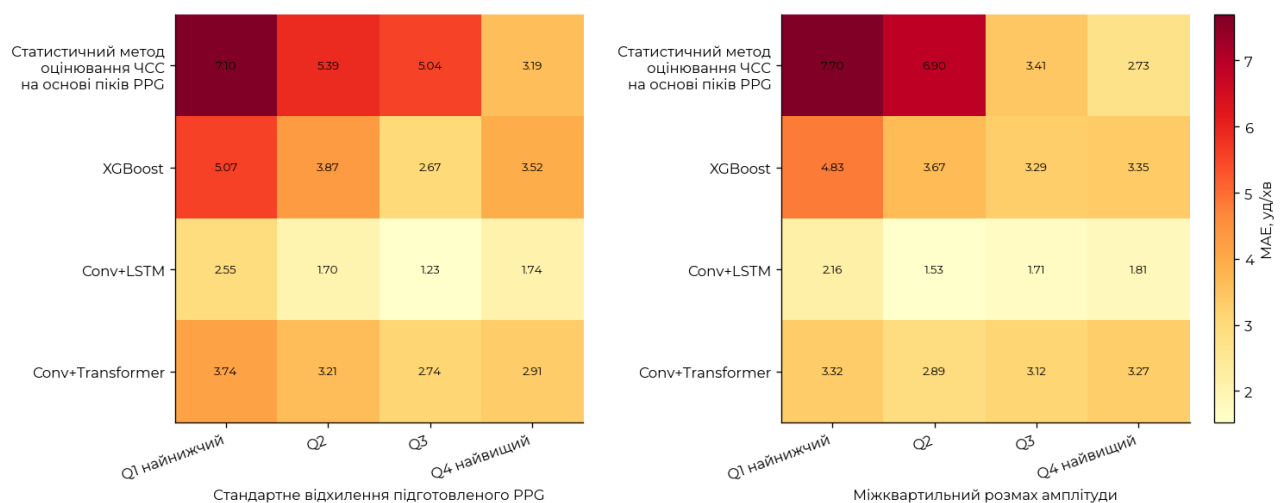


Рисунок 3.7 – Теплова матриця похибки за квантилями якості PPG

Аналіз за якістю сигналу підтверджує одну з головних переваг Conv+LSTM: модель краще зберігає точність у несприятливих умовах, ніж класичний піковий метод і таблична модель на ручних ознаках.

3.10. Аналіз кореляції помилок між моделями

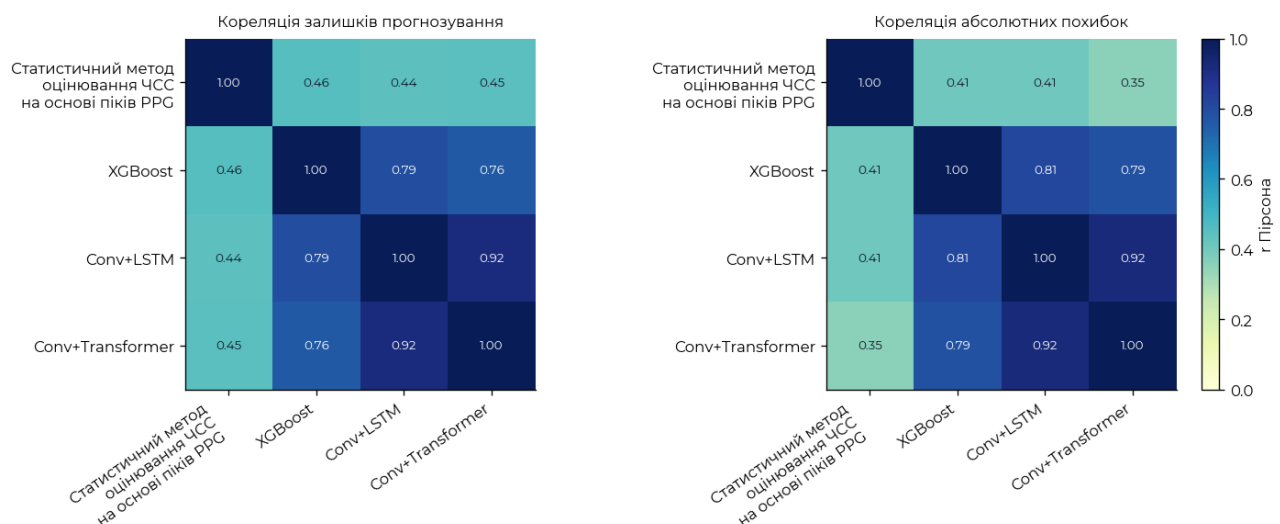
Залишки навчальних моделей є досить схожими. Найвища кореляція спостерігається між Conv+LSTM і Conv+Transformer – 0,92. Це означає, що ці дві моделі часто помиляються в одному напрямку: якщо одна занижує або завищує ЧСС, інша часто має схожу тенденцію.

Кореляція між XGBoost і Conv+LSTM становить 0,79, а між XGBoost і Conv+Transformer – 0,76. Це також свідчить про частково спільний характер помилок навчальних моделей.

Статистичний метод PPG має нижчу кореляцію з навчальними моделями, приблизно 0,44–0,46. Це означає, що природа його помилок відрізняється. Це очікувано, оскільки статистичний підхід залежить від пошуку піків, тоді як навчальні моделі використовують ширший набір патернів.

Кореляції абсолютних похибок є нижчими. Це означає, що навіть якщо моделі мають схожий напрямок помилки, великі помилки не завжди виникають на тих самих вікнах. Такий результат може бути корисним для майбутніх ансамблевих підходів або систем оцінювання невизначеності.

Матриці кореляції залишків прогнозування та абсолютних похибок між моделями наведено на рисунку 3.8. Кореляційний аналіз показує, що моделі частково вловлюють спільні закономірності, але не є повністю однаковими за характером помилок. Це означає, що в майбутньому можна досліджувати комбінування моделей або використання різних підходів для виявлення ненадійних прогнозів.



Рисunek 3.8 – Кореляція залишків прогнозування і абсолютних похибок між моделями

3.11. Оцінювання швидкодії та компактності моделей

Окрім точності, було оцінено компактність і швидкодію моделей. Результати наведено в таблиці 3.12.

Таблиця 3.12 – Швидкодія та компактність моделей

Модель	Представлення	Розмір	Кількість параметрів	Медіанний час прогнозу
Статистичний метод PPG	Пошук піків у 30-секундному PPG	0 КБ	0	0,065 мс/вікно
XGBoost	133 ручні ознаки	1172,5 КБ	Не застосовується	0,173 мс/вікно
Conv+LSTM	30-секундний PPG, 3750 відліків	237,0 КБ	57121	2,543 мс/вікно
Conv+Transformer	30-секундний PPG, 3750 відліків	252,1 КБ	59745	0,723 мс/вікно

Найшвидшим є статистичний метод, оскільки він не використовує навчальну модель. XGBoost також дуже швидкий, але для нього вказаний час відповідає прогнозу дерев після формування ознак. У практичному застосуванні потрібно додатково враховувати час обчислення ручних ознак.

Conv+Transformer має швидший прямий прохід, ніж Conv+LSTM, але поступається їй за точністю. Conv+LSTM є найповільнішою серед чотирьох моделей, але її швидкодія все одно значно перевищує вимоги реального часу. Один прогноз для 30-секундного вікна займає приблизно 2,543 мс, що відповідає близько 393 вікнам за секунду.

Обидві нейронні моделі є компактними: їхній розмір становить приблизно 237–252 КБ, а кількість параметрів – близько 60 тисяч. Це свідчить про можливість практичного використання таких моделей у системах, де важливі обмеження на пам'ять і швидкодію.

3.12. Аналіз результатів після фільтрації фізіологічно правдоподібного діапазону ЧСС

Оскільки у референсному каналі були виявлені екстремальні значення 0 і 300 уд/хв, було проведено додаткове оцінювання після фільтрації референсної ЧСС у діапазоні 30–220 уд/хв.

Результати повторного оцінювання після фільтрації референсної ЧСС у діапазоні 30–220 уд/хв наведено в таблиці 3.13. Після вилучення екстремальних референсних значень якості усіх моделей покращується, особливо RMSE. Для Conv+LSTM MAE зменшується з 1,803 до 1,641 уд/хв, а RMSE – з 6,075 до 4,941 уд/хв. R^2 зростає з 0,842 до 0,890. Це підтверджує, що значна частина великих похибок пов'язана з рідкісними або аномальними значеннями референсної ЧСС. Водночас навіть після такої фільтрації Conv+LSTM залишається найкращою моделлю за основними метриками.

Таблиця 3.13 – Результати після фільтрації ЧСС у діапазоні 30–220 уд/хв

Модель	MAE	RMSE	Медіанна AE	P95 AE	Зсув	R^2
Статистичний метод PPG	4,996	14,547	0,769	37,463	+3,656	0,045
XGBoost	3,627	6,741	1,857	12,738	-0,196	0,795
Conv+LSTM	1,641	4,941	0,491	6,265	-0,608	0,890
Conv+Transformer	3,000	5,732	1,968	8,297	-2,141	0,852

3.13. Узагальнення отриманих результатів і практичні обмеження

Проведені експерименти показали, що компактні нейронні моделі є ефективними для оцінювання частоти серцевих скорочень за PPG-сигналом. Найкращою серед розглянутих моделей стала Conv+LSTM. Її перевага

проявилася в усіх ключових напрямках: середній абсолютній похибці, медіанній похибці, 95-му процентилі абсолютної похибки, R^2 , порогових метриках і аналізі на рівні клінічних кейсів.

Перевага Conv+LSTM, імовірно, пояснюється поєднанням двох механізмів. Згортковий блок виділяє локальні морфологічні ознаки PPG-хвилі, а двонаправлений LSTM узагальнює ритмічну часову структуру всього 30-секундного вікна. На відміну від статистичного методу PPG, модель не залежить повністю від одного етапу пошуку піків. На відміну від XGBoost, вона працює з повною часовою формою сигналу, а не лише зі стиснутим набором ручних ознак.

Conv+Transformer також показала високу якість, але поступилася Conv+LSTM за MAE і мала помітніший від'ємний зсув. Це може свідчити про те, що для цієї конкретної задачі рекурентна агрегація після згорткового стискання краще відповідає ритмічній природі PPG-сигналу. Водночас швидший прямий прохід Conv+Transformer робить її цікавою для подальших досліджень.

XGBoost є сильним табличним підходом і має майже нульовий середній зсув, але поступається нейронним моделям за точністю. Його результати показують, що ручні ознаки дійсно містять значну інформацію про ЧСС, але не повністю передають часову структуру сигналу.

Статистичний метод PPG корисний як інтерпретований базовий підхід. Його низька медіанна похибка показує, що на чистих сигналах класичний пошук піків може бути дуже точним. Проте високі RMSE і P95 AE показують його нестабільність на складних вікнах.

Основні обмеження роботи такі:

- Референсний канал Solar8000/HR не є абсолютно безпомилковим і містить аномальні значення.
- Розподіл ЧСС є нерівномірним, більшість даних знаходиться в діапазоні 50–100 уд/хв.
- Високі й екстремальні значення ЧСС представлені мало, тому моделі гірше працюють у цих діапазонах.

- Оцінювання виконано на відкладених клінічних кейсах VitalDB, але не на зовнішньому незалежному наборі даних.
- У роботі не розв'язувалася задача клінічної валідації моделі для реального медичного застосування.

Попри ці обмеження, отримані результати підтверджують доцільність використання компактних нейронних моделей для оцінювання біометричних показників людини за фізіологічними сигналами.

3.14. Висновки до розділу 3

У третьому розділі було проведено експериментальне оцінювання розроблених підходів до автоматизованого оцінювання частоти серцевих скорочень за фотоплетизмографічним сигналом. Для цього використано сформований набір даних на основі клінічних записів VitalDB, який містив 6157 клінічних кейсів і 2129005 неперекривних 30-секундних вікон. Фінальне порівняння моделей виконано на 318268 спільних тестових вікнах після фільтрації за якістю сигналу та вилучення прикладів із нечисловими прогнозами.

Основне порівняння показало, що найкращі результати отримала модель Conv+LSTM. Вона досягла MAE 1,803 уд/хв, RMSE 6,075 уд/хв, R^2 0,842, медіанної абсолютної похибки 0,493 уд/хв і P95 AE 6,579 уд/хв. Частка прогнозів із похибкою не більше 5 уд/хв становила 93,64%, а з похибкою не більше 10 уд/хв — 96,59%. Це підтверджує, що модель забезпечує вищу точність і стійкість оцінювання ЧСС порівняно з класичним статистичним методом.

Модель Conv+Transformer також показала високу якість і перевершила XGBoost за середньою абсолютною похибкою, однак поступилася Conv+LSTM і мала помітніший від'ємний зсув. Модель XGBoost продемонструвала майже нульовий середній зсув, але за більшістю метрик поступилася нейронним моделям. Класичний статистичний метод PPG мав низьку медіанну абсолютну похибку, проте характеризувався великим 95-м перцентилем абсолютної похибки та найгіршим RMSE, що свідчить про нестабільність на складних або зашумлених вікнах.

Аналіз на рівні клінічних кейсів підтвердив перевагу Conv+LSTM не лише на рівні окремих вікон, а й на рівні окремих клінічних записів. Для цієї моделі медіана MAE на рівні кейсів становила 1,142 уд/хв, середнє значення — 1,937 уд/хв, а 95-й перцентиль — 5,885 уд/хв. Це свідчить про стабільність моделі при переході до нових клінічних кейсів.

Аналіз Бленда-Альтмана показав, що Conv+LSTM має найвужчі межі узгодження з референсною ЧСС і невеликий середній зсув -0,448 уд/хв. Калібраційний аналіз підтвердив, що модель найкраще відтворює центральний діапазон ЧСС, у якому зосереджена основна частина даних. Водночас аналіз похибки за діапазонами ЧСС показав, що на високих та екстремальних значеннях похибка всіх моделей зростає через малу кількість таких прикладів, тенденцію моделей до заниження високих значень і можливі артефакти референсного каналу.

Оцінювання впливу якості PPG-сигналу показало, що низька амплітуда або низька варіативність підготовленого сигналу погіршують результати всіх моделей. Найбільш чутливим до якості сигналу виявився статистичний метод на основі піків, тоді як Conv+LSTM продемонструвала найстабільнішу поведінку за квантилями показників якості PPG.

Оцінювання швидкодії та компактності показало, що нейронні моделі залишаються достатньо компактними для практичного використання. Conv+LSTM має 57121 навчальний параметр, розмір приблизно 237 КБ і медіанний час прогнозу 2,543 мс на одне 30-секундне вікно. Conv+Transformer має 59745 параметрів, розмір приблизно 252,1 КБ і медіанний час прогнозу 0,723 мс. Отже, обидві моделі мають достатню швидкодію для подальшого використання в програмних модулях систем моніторингу стану людини.

Додаткове оцінювання після фільтрації референсної ЧСС у фізіологічно правдоподібному діапазоні 30–220 уд/хв показало покращення якості всіх моделей. Для Conv+LSTM MAE зменшилася з 1,803 до 1,641 уд/хв, RMSE — з 6,075 до 4,941 уд/хв, а R^2 зросло з 0,842 до 0,890. Це підтверджує, що частина

великих похибок пов'язана з екстремальними або артефактними значеннями референсного каналу.

Таким чином, результати третього розділу підтверджують досягнення мети кваліфікаційної роботи: створений програмний підхід на основі моделі Conv+LSTM забезпечив підвищення точності та стійкості автоматизованого оцінювання частоти серцевих скорочень за PPG-сигналом порівняно з класичним статистичним методом і може бути використаний як основа для подальшого розвитку програмних модулів систем моніторингу біометричних показників людини.

ВИСНОВКИ

У кваліфікаційній роботі розв’язано задачу оцінювання біометричного показника людини – частоти серцевих скорочень – за фотоплетизмографічним сигналом із використанням методів машинного та глибокого навчання. У межах теми «Оцінка біометричних показників людини з використанням глибокого навчання» було обґрунтовано вибір частоти серцевих скорочень як цільового біометричного показника, а фотоплетизмографічного сигналу – як вхідного джерела інформації.

У першому розділі було проаналізовано предметну область оцінювання біометричних показників людини за фізіологічними сигналами. Розглянуто частоту серцевих скорочень як фізіологічний показник, природу фотоплетизмографічного сигналу, класичні методи оцінювання ЧСС за піками PPG, підходи машинного навчання на основі ручних ознак і моделі глибокого навчання для одновимірних часових рядів. Проведений аналіз показав, що класичні методи є простими й інтерпретованими, але недостатньо стабільними на шумних або деформованих сигналах. Це обґрунтовує доцільність використання навчальних моделей, зокрема компактних нейронних архітектур.

У другому розділі було сформульовано постановку задачі, визначено об’єкт, предмет, мету й завдання роботи, описано обмеження дослідження та розроблену методику. Було сформовано набір даних на основі клінічних записів VitalDB. Вхідним сигналом обрано канал SNUADC/PLETH, а референсним значенням ЧСС – канал Solar8000/HR. Загалом було використано 6157 клінічних кейсів і сформовано 2129005 неперекривних 30-секундних вікон. Розбиття на навчальну, валідаційну й тестову вибірки виконано за ідентифікатором клінічного кейсу, що запобігає витоку даних між вибірками.

Було реалізовано попередню обробку PPG-сигналу, яка включає заповнення пропусків, смугову фільтрацію, зменшення частоти дискретизації з 500 Гц до 125 Гц і стійку нормалізацію. Для кожного вікна розраховано показники якості сигналу: частку валідних відліків PPG-сигналу,

міжквартильний розмах амплітуди PPG-сигналу, стандартне відхилення підготовленого PPG-сигналу та частку потенційно обмежених за амплітудою значень. Ці показники використовувалися для фільтрації, побудови ознак та аналізу впливу якості сигналу на похибку.

У роботі реалізовано та порівняно чотири підходи до оцінювання ЧСС. Перший – статистичний метод PPG, тобто статистичний метод на основі пошуку піків PPG і медіанного міжпікового інтервалу. Другий – XGBoost на основі 133 ручних ознак, які описують якість, статистичні, часові та спектральні властивості сигналу. Третій – Conv+LSTM, компактна нейронна модель, що поєднує згорткові шари та двонапрямлений LSTM. Четвертий – Conv+Transformer, компактна модель, що поєднує згорткове перетворення сигналу та кодувальник Transformer.

У третьому розділі проведено експериментальне оцінювання моделей. Фінальне порівняння виконано на 318268 спільних тестових вікнах після фільтрації за якістю сигналу. Найкращі результати отримала модель Conv+LSTM. Вона досягла MAE 1,803 уд/хв, RMSE 6,075 уд/хв, R2 0,842, медіанної абсолютної похибки 0,493 уд/хв і P95 AE 6,579 уд/хв. Частка прогнозів із похибкою не більше 5 уд/хв становила 93,64%, а з похибкою не більше 10 уд/хв – 96,59%.

Conv+Transformer посіла друге місце за середньою абсолютною похибкою: MAE становила 3,151 уд/хв. XGBoost мав MAE 3,782 уд/хв і майже нульовий середній зсув, але поступився нейронним моделям за більшістю метрик. Статистичний метод PPG мав MAE 5,183 уд/хв, низьку медіанну абсолютну похибку 0,770 уд/хв, але дуже високий P95 AE 38,940 уд/хв і RMSE 15,164 уд/хв. Це підтверджує, що піковий метод часто працює точно на чистих сигналах, але має довгий хвіст великих помилок.

Аналіз на рівні клінічних кейсів підтвердив перевагу Conv+LSTM. Медіана MAE на рівні клінічних кейсів для цієї моделі становила 1,142 уд/хв, середня MAE на рівні клінічних кейсів – 1,937 уд/хв, а 95-й перцентиль MAE на рівні

клінічних кейсів – 5,885 уд/хв. Це означає, що перевага моделі зберігається не лише на великій кількості окремих вікон, а й на рівні окремих клінічних записів.

Аналіз Бленда-Альтмана показав, що Conv+LSTM має найвужчі межі узгодження та невеликий середній зсув -0,448 уд/хв. XGBoost мав майже нульовий зсув, але ширші межі узгодження. Conv+Transformer мав помітніший від’ємний зсув -1,985 уд/хв, а статистичний метод – додатний зсув +3,841 уд/хв і найширші межі узгодження.

Аналіз похибки за діапазонами ЧСС показав, що Conv+LSTM найкраще працює в центральному діапазоні 50–90 уд/хв, де зосереджена більшість даних. У цьому діапазоні її MAE часто становить близько 1–2 уд/хв. Водночас у високих і екстремальних діапазонах ЧСС похибка всіх моделей зростає. Це пов’язано з малою кількістю таких прикладів, тенденцією моделей до заниження високих значень і наявністю потенційних артефактів у референсному каналі.

Аналіз якості PPG-сигналу показав, що низька амплітуда або низька варіативність підготовленого сигналу погіршують результати всіх моделей. Статистичний метод виявився найбільш чутливим до якості сигналу, оскільки його робота прямо залежить від правильного пошуку піків. Conv+LSTM була найстабільнішою за квартилями якості сигналу, що є одним із важливих практичних результатів роботи.

Оцінювання швидкодії та компактності показало, що всі моделі працюють швидше за вимоги реального часу. Conv+LSTM має 57121 навчальний параметр, розмір приблизно 237 КБ і медіанний час прогнозу 2,543 мс на одне 30-секундне вікно. Conv+Transformer має 59745 параметрів, розмір приблизно 252,1 КБ і медіанний час прогнозу 0,723 мс. Це підтверджує, що компактні нейронні моделі можуть бути практично придатними для задач оцінювання ЧСС за PPG.

Додатковий аналіз після фільтрації референсної ЧСС у фізіологічно правдоподібному діапазоні 30–220 уд/хв показав покращення якості всіх моделей. Для Conv+LSTM MAE зменшилася з 1,803 до 1,641 уд/хв, RMSE – з 6,075 до 4,941 уд/хв, а R^2 зросло з 0,842 до 0,890. Це підтверджує, що частина

великих похибок пов'язана з екстремальними або артефактними значеннями референсного каналу.

Основним результатом роботи є розроблений та експериментально перевірений підхід до оцінювання частоти серцевих скорочень за PPG-сигналом. Найкраща модель Conv+LSTM забезпечила високу точність, стабільність на рівні клінічних кейсів, компактність і достатню швидкодію. Отримані результати підтверджують доцільність використання глибокого навчання для оцінювання біометричних показників людини за фізіологічними сигналами.

Основні напрями подальшого розвитку роботи включають очищення або уточнення референсних міток, використання ЧСС, отриманої з ECG-сигналу, як додаткового або альтернативного еталона, підсилення навчання на високих значеннях ЧСС, застосування методів оцінювання невизначеності, реалізацію механізму відмови від прогнозу на низькоякісних вікнах, зовнішню перевірку на незалежних наборах даних і оптимізацію моделей для практичного розгортання.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Наказ Харківського національного університету імені В. Н. Каразіна про внесення змін до наказу від 29.01.2026 № 4101-5/225 «Про затвердження тем кваліфікаційних робіт». Харківський національний університет імені В. Н. Каразіна, 2026.
2. Терміни підготовки кваліфікаційних робіт до захисту. Комісія № 157. Галузь знань 12 Інформаційні технології. Спеціальність 122 Комп'ютерні науки. Ступінь бакалавра: освітня програма «Комп'ютерні науки». Харківський національний університет імені В. Н. Каразіна, 2026.
3. Методичні рекомендації щодо виконання кваліфікаційних робіт студентами спеціальності 122 Комп'ютерні науки першого (бакалаврського) рівня вищої освіти / уклад. М. В. Ткачук, Д. В. Зінов'єв, Н. Г. Хацько, Н. О. Мусс. Харків : ХНУ імені В. Н. Каразіна, 2020. 43 с.
4. Кваліфікаційна робота : методичні вказівки з написання та оформлення для здобувачів вищої освіти [Електронне видання] / уклад. Б. В. Самородов, О. М. Головка, І. К. Журавльова, С. В. Єльцов. Харків : ХНУ імені В. Н. Каразіна, 2026. 12 с.
5. Порядок проведення перевірки кваліфікаційних робіт, наукових праць та навчальних видань щодо наявності запозичень з інших документів : затверджено рішенням Вченої ради Харківського національного університету імені В. Н. Каразіна від 24.02.2020 р., протокол № 5; введено в дію наказом ректора № 0204-1/088 від 27.02.2020 р. Харків : ХНУ імені В. Н. Каразіна, 2020.
6. Lee H.-C., Park Y., Yoon S. B., Yang S. M., Park D., Jung C.-W. VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients. Scientific Data. 2022. Vol. 9. Article 279. DOI: 10.1038/s41597-022-01411-5.
7. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement. 2007. Vol. 28, No. 3. P. R1–R39. DOI: 10.1088/0967-3334/28/3/R01.

8. Elgendi M. On the analysis of fingertip photoplethysmogram signals. *Current Cardiology Reviews*. 2012. Vol. 8, No. 1. P. 14–25. DOI: 10.2174/157340312801215782.
9. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. Cambridge, MA : MIT Press, 2016. 800 p.
10. Bishop C. M. *Pattern Recognition and Machine Learning*. New York : Springer, 2006. 738 p.
11. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York : Springer, 2009. 745 p. DOI: 10.1007/978-0-387-84858-7.
12. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, USA, 2016. P. 785–794. DOI: 10.1145/2939672.2939785.
13. Hochreiter S., Schmidhuber J. Long Short-Term Memory. *Neural Computation*. 1997. Vol. 9, No. 8. P. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
14. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention Is All You Need. *Advances in Neural Information Processing Systems*. 2017. Vol. 30.
15. Bland J. M., Altman D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986. Vol. 327, No. 8476. P. 307–310. DOI: 10.1016/S0140-6736(86)90837-8.
16. SciPy Documentation. Signal processing package. URL: <https://docs.scipy.org/doc/scipy/reference/signal.html> (дата звернення: 25.05.2026).
17. PyTorch Documentation. Deep learning framework documentation. URL: <https://docs.pytorch.org/docs/stable/index.html> (дата звернення: 25.05.2026).
18. XGBoost Documentation. Scalable gradient boosting library documentation. URL: <https://xgboost.readthedocs.io/> (дата звернення: 25.05.2026).
19. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V. et al. *Scikit-learn: Machine*

Learning in Python. *Journal of Machine Learning Research*. 2011. Vol. 12. P. 2825–2830.

20. Harris C. R., Millman K. J., van der Walt S. J., Gommers R., Virtanen P., Cournapeau D., Wieser E., Taylor J., Berg S., Smith N. J. et al. Array programming with NumPy. *Nature*. 2020. Vol. 585. P. 357–362. DOI: 10.1038/s41586-020-2649-2.

ДОДАТОК А. ФРАГМЕНТИ ПРОГРАМНОЇ РЕАЛІЗАЦІЇ

Лістинг А.1 – Загальна логіка формування 30-секундних вікон

```
def extract_case_windows(case_id: int, split_name: str, config:
dict):
    raw_fs_hz = int(config["raw_fs_hz"])
    window_sec = int(config["window_sec"])
    step_sec = int(config["step_sec"])
    raw_samples = int(config["raw_samples"])
    model_samples = int(config["model_samples"])

    ppg_raw = load_ppg_track(case_id, config)
    ref_hr = load_reference_hr(case_id, config)
    ppg_prepared_full = preprocess_ppg(ppg_raw, config)

    duration_sec = min(len(ppg_raw) / raw_fs_hz, len(ref_hr))
    rows, raw_windows, model_windows = [], [], []

    window_starts = np.arange(0, duration_sec - window_sec +
1e-9, step_sec)

    for window_id, start_sec in enumerate(window_starts):
        end_sec = start_sec + window_sec

        start_idx = int(round(start_sec * raw_fs_hz))
        end_idx = int(round(end_sec * raw_fs_hz))

        raw_window = ppg_raw[start_idx:end_idx]
        prepared_window_raw_fs =
ppg_prepared_full[start_idx:end_idx]

        ref_slice = ref_hr.iloc[int(start_sec):int(end_sec)]
        label_ref_hr_bpm = float(ref_slice.mean())

        if not np.isfinite(label_ref_hr_bpm):
```

```

        continue

    if len(raw_window) != raw_samples:
        continue

    model_window =
downsample_window(prepared_window_raw_fs, config)
    if len(model_window) != model_samples:
        continue

    model_window = robust_zscore(model_window)

    valid_ratio,          amplitude_iqr,          prepared_std,
clipped_ratio = (
        ppg_window_quality(raw_window,
prepared_window_raw_fs)
    )

    rows.append({
        "split": split_name,
        "case_id": int(case_id),
        "window_id": int(window_id),
        "start_sec": float(start_sec),
        "end_sec": float(end_sec),
        "window_mid_sec": float(start_sec + window_sec /
2),

        "label_ref_hr_bpm": label_ref_hr_bpm,
        "ppg_valid_ratio": valid_ratio,
        "ppg_amplitude_iqr": amplitude_iqr,
        "ppg_prepared_std": prepared_std,
        "ppg_clipped_ratio": clipped_ratio,
    })

    raw_windows.append(raw_window.astype(np.float32))
    model_windows.append(model_window.astype(np.float32))

return rows, raw_windows, model_windows

```

Лістинг А.2 – Попередня обробка PPG-сигналу

```

def fill_missing_linear_both(values) -> np.ndarray:
    filled = np.asarray(values, dtype=np.float64).copy()
    finite = np.isfinite(filled)

    if finite.all() or not finite.any():
        return filled

    sample_indices = np.arange(len(filled))
    filled[~finite] = np.interp(
        sample_indices[~finite],
        sample_indices[finite],
        filled[finite],
    )
    return filled

def preprocess_ppg(ppg_raw, config: dict) -> np.ndarray:
    filled = fill_missing_linear_both(ppg_raw)

    fs_hz = int(config["raw_fs_hz"])
    nyquist = fs_hz / 2
    low_hz = float(config["bandpass_low_hz"])
    high_hz = float(config["bandpass_high_hz"])
    order = int(config["filter_order"])

    sos = signal.butter(
        order,
        [low_hz / nyquist, high_hz / nyquist],
        btype="bandpass",
        output="sos",
    )

    filtered = signal.sosfiltfilt(sos, filled)

```

```

return filtered.astype(np.float32)

def downsample_window(values, config: dict) -> np.ndarray:
    raw_fs_hz = int(config["raw_fs_hz"])
    model_fs_hz = int(config["model_fs_hz"])

    target_len = int(round(len(values) * model_fs_hz /
raw_fs_hz))
    downsampled = signal.resample_poly(
        values,
        up=model_fs_hz,
        down=raw_fs_hz,
    )

    return downsampled[:target_len].astype(np.float32)

def robust_zscore(values, eps=1e-8) -> np.ndarray:
    values = np.asarray(values, dtype=np.float64)

    median = np.nanmedian(values)
    mad = np.nanmedian(np.abs(values - median))
    scale = 1.4826 * mad

    if not np.isfinite(scale) or scale < eps:
        scale = np.nanstd(values)

    if not np.isfinite(scale) or scale < eps:
        return np.zeros_like(values, dtype=np.float32)

    return ((values - median) / scale).astype(np.float32)

```

Лістинг А.3 – Обчислення показників якості PPG-вікна

```

def ppg_window_quality(raw_window, prepared_window):
    valid_ratio = float(np.isfinite(raw_window).mean())

```

```

finite_raw = raw_window[np.isfinite(raw_window)]
if finite_raw.size == 0:
    return valid_ratio, np.nan, np.nan, np.nan

amplitude_iqr = float(
    np.subtract(*np.nanpercentile(finite_raw, [75, 25]))
)

prepared_std = float(np.nanstd(prepared_window))

p_low, p_high = np.nanpercentile(finite_raw, [0.5, 99.5])
clipped_ratio = float(
    np.mean((finite_raw <= p_low) | (finite_raw >= p_high))
)

return valid_ratio, amplitude_iqr, prepared_std,
clipped_ratio

```

Лістинг A.4 – Архітектура Conv+LSTM

```

class ConvLSTMRegressor(nn.Module):
    def __init__(self, hidden_size=48, lstm_layers=1,
dropout=0.15):
        super().__init__()

        self.stem = nn.Sequential(
            nn.Conv1d(1, 16, kernel_size=9, stride=2,
padding=4, bias=False),
            nn.BatchNorm1d(16),
            nn.SiLU(),

            nn.Conv1d(16, 24, kernel_size=7, stride=2,
padding=3, bias=False),
            nn.BatchNorm1d(24),
            nn.SiLU(),

```

```

        nn.Conv1d(24,      32,      kernel_size=5,      stride=2,
padding=2, bias=False),
        nn.BatchNorm1d(32),
        nn.SiLU(),

        nn.Conv1d(32,      48,      kernel_size=5,      stride=2,
padding=2, bias=False),
        nn.BatchNorm1d(48),
        nn.SiLU(),
        nn.Dropout(dropout),
    )

    self.lstm = nn.LSTM(
        input_size=48,
        hidden_size=hidden_size,
        num_layers=lstm_layers,
        batch_first=True,
        bidirectional=True,
    )

    self.head = nn.Sequential(
        nn.LayerNorm(hidden_size * 2),
        nn.Linear(hidden_size * 2, 48),
        nn.SiLU(),
        nn.Dropout(dropout),
        nn.Linear(48, 1),
    )

    def forward(self, x):
        features = self.stem(x).transpose(1, 2)
        encoded, _ = self.lstm(features)
        pooled = encoded.mean(dim=1)
        return self.head(pooled).squeeze(-1)

```

Лістинг А.5 – Обчислення основних метрик

```

def regression_metrics(y_true: np.ndarray, y_pred: np.ndarray)
-> pd.Series:
    y_true = np.asarray(y_true, dtype=float)
    y_pred = np.asarray(y_pred, dtype=float)

    mask = np.isfinite(y_true) & np.isfinite(y_pred)
    y_true = y_true[mask]
    y_pred = y_pred[mask]

    error = y_pred - y_true
    abs_error = np.abs(error)

    return pd.Series({
        "n_windows": int(len(y_true)),
        "MAE, bpm": float(np.mean(abs_error)),
        "RMSE, bpm": float(np.sqrt(np.mean(error ** 2))),
        "Median AE, bpm": float(np.median(abs_error)),
        "P90 AE, bpm": float(np.percentile(abs_error, 90)),
        "P95 AE, bpm": float(np.percentile(abs_error, 95)),
        "Bias, bpm": float(np.mean(error)),
        "SD error, bpm": float(np.std(error, ddof=1)),
        "R2": float(r2_score(y_true, y_pred)),
        "AE <= 3 bpm, %": float(100 * np.mean(abs_error <= 3)),
        "AE <= 5 bpm, %": float(100 * np.mean(abs_error <= 5)),
        "AE <= 10 bpm, %": float(100 * np.mean(abs_error <=
10))),
    })

```