Issue 18, part 2, 2025

https://doi.org/10.23939/sisn2025.18.2.113

УДК 004.85:616.12-073.97:519.254

МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ КЛАСИФІКАЦІЇ ЕЛЕКТРОКАРДІОСИГНАЛІВ ЗА РИТМІЧНИМИ ТА МОРФОЛОГІЧНИМИ ОЗНАКАМИ

Любомир Мосій[1], Андрій Сверстюк[2]

[1] Тернопільський національний технічний університет імені Івана Пулюя,
кафедра комп'ютерних наук, Тернопіль, Україна
[2] Тернопільський національний медичний університет імені І.Я. Горбачевського,
кафедра медичної інформатики, Тернопіль, Україна
E-mail: lmosiy@gmail.com, ORCID: 0009-0000-9778-331X
E-mail: sverstyuk@tdmu.edu.ua, ORCID: 0000-0001-8644-0776

© Мосій Л.Є., Сверстюк А.С., 2025

У статті представлено експериментальне дослідження ефективності методів машинного навчання для класифікації електрокардіосигналів за ритмічними та морфологічними ознаками з використанням інформаційної технології на основі математичного апарату циклічних випадкових процесів. Розглянуто проблему автоматизованого виявлення передсердних аритмій, зокрема фібриляції та тріпотіння передсердь, які характеризуються комплексними змінами як у морфології зубців ЕКС, так і в часових інтервалах серцевих циклів.

Для класифікації патологічних станів досліджено чотири алгоритми виявлення аномалій: OneClassSVM з радіально-базисною функцією, IsolationForest, Local Outlier Factor (LOF) та EllipticEnvelope. Проаналізовано вплив методів попередньої обробки даних (StandardScaler для стандартизації та PCA для зниження розмірності) на точність класифікації. Експериментальні результати показали, що застосування попередньої обробки критично важливе для класифікації морфологічних порушень, підвищуючи точність з 50-83 % до 100 % для фібриляції передсердь. Алгоритм LOF продемонстрував найбільш стабільні результати (83-100 %) для різних типів патологій. При класифікації ритмічних порушень методи IsolationForest, LOF та EllipticEnvelope показали однаково високу ефективність (89 %), при цьому попередня обробка не призвела до суттєвого покращення результатів.

Ключові слова – моделювання електрокардіосигналів, інформаційна технологія, циклічний дискретний випадковий процес, амплітудно-часові характеристики, часова функція ритму, амплітудна варіабельність, класифікація сигналів, штучний інтелект (AI), система машинного навчання (MLS).

Постановка проблеми

Автоматизований аналіз електрокардіосигналів (ЕКС) залишається критично важливим завданням сучасної медичної діагностики, особливо в контексті виявлення передсердних аритмій, які часто мають субклінічний перебіг та можуть призводити до серйозних ускладнень. Незважаючи на значні досягнення в області машинного навчання для аналізу ЕКС, існує ряд невирішених проблем, які обмежують ефективність існуючих підходів.

Багато сучасних методів класифікації ЕКС зосереджені або на аналізі морфологічних характеристик окремих комплексів, або на оцінці ритмічних параметрів, не забезпечуючи комплексного підходу до одночасного врахування обох типів ознак. Це призводить до неповної характеристики патологічних станів, оскільки такі аритмії як фібриляція та тріпотіння передсердь проявляються змінами як у морфології зубців ЕКС, так і в часових інтервалах між серцевими циклами.

Таким чином, актуальною науковою проблемою є розроблення та валідація комплексного підходу до класифікації ЕКС, який би інтегрував аналіз як морфологічних, так і ритмічних ознак на основі математичного апарату циклічних випадкових процесів (ЦВП), з оптимальним вибором методів попередньої обробки даних та алгоритмів машинного навчання для різних типів передсердних аритмій.

Аналіз останніх досліджень та публікацій

Машинне навчання (Machine learning – ML) революціонізувало аналіз електрокардіосигналів (ЕКС), забезпечивши автоматизовану інтерпретацію з підвищеною швидкістю, точністю та масштабованістю. Моделі ML можуть вивчати закономірності безпосередньо з необроблених або попередньо оброблених ЕКС, що дозволяє їм виявляти такі особливості, які часто пропускаються при ручному аналізі. Ця зміна зменшує залежність від досвідчених клініцистів і розширює доступність, що особливо важливо в умовах обмеженої кардіологічної експертизи (Wasimuddin, 2021).

Переваги методів ML у порівнянні з ручними підходами особливо очевидні з точки зору швидкості та послідовності. Алгоритми можуть швидко опрацьовувати великі обсяги даних ЕКС, що полегшує моніторинг майже в режимі реального часу та ранню діагностику. Більше того, моделі ML демонструють відтворюваність, мінімізуючи міжспостерігацьку варіативність, яка є типовою для людської інтерпретації. Сучасні тенденції відображають прогресивний перехід від традиційних методів опрацювання сигналів до передових методологій ML та глибинного навчання (Deep Learning – DL). Класичні методи зосереджуються на явній розробці ознак та класифікаторах на основі правил, тоді як підходи DL пропонують комплексні рішення, здатні до автоматичного вилучення ознак та ієрархічного навчання представлення на основі складних наборів даних (Kumar, 2023). Ці досягнення збагатили аналітичні можливості ЕКС, розширивши потенціал прогнозної діагностики, персоналізованої медицини та інтеграції з носимими медичними пристроями (Gupta, 2023).

У цьому огляді розглядаються методи машинного навчання, що застосовуються для класифікації ЕКС, та надається всебічний аналіз традиційних підходів, архітектур глибинного навчання та гібридних стратегій. Методологія огляду охоплює рецензовані публікації за період 2018–2025 років, зосереджуючись на методах, що пройшли клінічну валідацію та мають опубліковані показники ефективності на стандартизованих наборах даних. Структура огляду починається з традиційних методів машинного навчання, продовжується підходами глибинного навчання та техніками ансамблю, а завершується оцінкою ефективності та майбутніми напрямками розвитку.

Традиційні методи машинного навчання, що застосовуються до класифікації ЕКС

Машини опорних векторів та їхні варіанти

Машини опорних векторів (Support Vector Machines – SVM) широко застосовуються в класифікації ЕКС завдяки їхній міцній теоретичній основі та ефективності в опрацюванні високорозмірних даних. Варіанти, такі як SVM з оптимізацією стохастичного градієнтного спуску (SVM-SGD) та класифікація опорних векторів (SVM-SVC), розширюють класичну структуру SVM, покращуючи обчислювальну ефективність та підвищуючи продуктивність класифікації за допомогою використання ядра (Yildirim, 2024).

Основні функції, включаючи лінійні, радіальні базисні функції (RBF) та поліноміальні ядра, дозволяють SVM проєктувати нелінійні дані ЕКС у вищі розмірні простори ознак, де класи стають роздільними. Ця здатність є особливо важливою для складної та нелінійної природи хвильових форм ЕКС. У кількох дослідженнях було застосовано варіанти моделей SVM для ефективного виявлення аритмії та інших серцево-судинних захворювань, використовуючи їх стійкість до перенавчання та здатність балансувати точність і відтворюваність (Rath, 2022). Ядро RBF продемонструвало чудову ефективність у фіксуванні нелінійних патернів ЕКС, досягаючи точності до 96 % у завданнях виявлення аритмії.

SVM, інтегрована з вибором ознак та оптимізованими параметрами ядра, продемонструвала високу чутливість у розрізненні передчасних та фібриляційних ударів, що виявилося особливо корисним у виявленні фібриляції передсердь та діагностиці ішемії. Гібридні моделі, що включають SVM, ще більше підвищують ефективність, використовуючи переваги як методів ядра, так і ансамблевого навчання для роботи з незбалансованими наборами даних (Yue, 2021). Однак обчислювальна складність, що масштабується з $O(n^2)$ для навчання, створює проблеми для додатків, що працюють у режимі реального часу, що вимагає використання наближених методів ядра або ієрархічних архітектур (Agrawal, 2022).

Дерева рішень та ансамблеві методи

Дерева рішень є інтуїтивним та інтерпретованим підходом до машинного навчання, що моделює рішення на основі критеріїв розділення для класифікації ЕКС на окремі серцеві стани. Однак окремі дерева часто страждають від надмірного пристосування та обмеженої генералізації. Ансамблеві методи, такі як випадкові ліси та дерева рішень з градієнтним підсиленням (Gradient Boosted Decision Trees – GBDT), вирішують ці проблеми шляхом побудови декількох дерев та агрегації їхніх прогнозів за допомогою голосування більшістю голосів або зваженого усереднення.

Ансамблеві моделі на основі дерев набули популярності в завданнях класифікації ЕКС завдяки своїй високій точності та надійності, особливо при опрацюванні шумних і незбалансованих даних. AdaBoost, техніка ансамблевого підсилення, адаптивно фокусується на прикладах, які важко класифікувати, покращуючи розпізнавання меншості класів у класифікації аритмій із підвищенням чутливості на 13 % для рідкісних аритмій. Підходи баґінгу, які створюють різноманітні набори класифікаторів за допомогою випадкового відбору, ще більше підвищують надійність і зменшують дисперсію (Śmigiel, 2024).

Порівняльний аналіз показує, що алгоритми градієнтного підсилення перевершують традиційні дерева рішень завдяки своєму механізму послідовного навчання, який ітеративно виправляє помилки. Ці моделі демонструють відмінні показники ефективності, такі як високі значення площі під кривою ROC (AUC) 0,92-0,97 та F1-бали, що перевищують 0,95, особливо в поєднанні з інженерією ознак та зменшенням розмірності. Підхід на основі ансамблю дерев підходить для використання в умовах обмежених ресурсів, забезпечуючи баланс між прогнозною здатністю та обчислювальними витратами (Hassaballah, 2023).

Моделі K-найближчих сусідів та Naive Bayes моделі

Класифікатори K-найближчих сусідів (K-NN) і наївних байєсів забезпечують обчислювально ефективні методи класифікації ЕКС, спираючись відповідно на метрики відстані та припущення про ймовірнісну незалежність. Простий механізм K-NN передбачає ідентифікацію найближчих мічених екземплярів у просторі ознак для класифікації невідомих зразків, що робить його привабливим для досліджень на ранніх стадіях або досліджень здійсненності (Ullah, 2024).

Моделі Naive Bayes припускають незалежність ознак і обчислюють апостеріорні ймовірності для класифікації, що робить їх обчислювально ефективними і особливо корисними для шумних або неповних даних ЕКС. Обидва методи використовуються для виявлення стадій сну та інших фізіологічних станів на основі витягнутих ознак варіабельності серцевого ритму з одноканального ЕКС (Coronado-Reyes, 2025). Останні тести на наборі даних MIT-BIH показують, що K-NN досягає точності 86,7 % проти 81,3 % для Naive Bayes у класифікації аритмії п’яти класів.

Обмеження виникають у високорозмірних наборах даних ЕКС, де обчислювальні витрати K-NN збільшуються, а метрики відстані стають менш дискримінаційними. Аналогічно, припущення про незалежність Naive Bayes часто не справджується для корельованих ознак ЕКС, що призводить до неоптимальної ефективності класифікації. Ретельний вибір ознак і зменшення розмірності мають вирішальне значення для подолання цих проблем, але в цілому ці моделі, як правило, працюють гірше, ніж більш досконалі алгоритми, з розривом у точності 8−15 % порівняно з підходами DL (Abdul Razak, 2023).

Підходи DL в аналізі ЕКС

Складні нейронні мережі для класифікації ЕКС

Складні нейронні мережі (Convolutional Neural Networks – CNN) трансформували аналіз ЕКС, забезпечуючи автоматичне вилучення ознак через ієрархічні складні шари, усуваючи необхідність ручного інжинірингу ознак. Архітектури CNN були адаптовані як для одномірних неопрацьованих вхідних даних ЕКС, так і для двовимірних представлень, таких як спектрограми або зображення, трансформовані на основі FFT (Kumar M., 2022).

Дослідження, в яких використовується FFT для перетворення ЕКС перед класифікацією CNN, демонструють поліпшення точності виявлення аритмії, причому моделі на основі CNN перевершують традиційні класифікатори на 3−7 %. Це пояснюється здатністю CNN ефективно фіксувати локалізовані сигнали та морфологічні нюанси форм хвиль ЕКС. Типові одномірні архітектури CNN використовують 4−8 конволюційних шарів з розмірами ядра 3−7 зразків, досягаючи 98,5 % точності при класифікації аритмії MIT-BIH (Ullah, 2024).

Рекурентні нейронні мережі LSTM, bi-LSTM та GRU

Рекурентні нейронні мережі (Recurrent Neural Networks – RNN), особливо мережі з довгою короткочасною пам'яттю (LSTM) та їх двонаправлені варіанти (bi-LSTM), призначені для моделювання часових залежностей у послідовних даних, таких як часові ряди ЕКС. За рахунок збереження внутрішніх станів пам'яті ці мережі ефективно фіксують динамічну поведінку серцевих ритмів у часі.

Bi-LSTM покращують контекстуальне розуміння шляхом опрацювання сигналів як у прямому, так і в зворотному часовому напрямку, що є особливо корисним для окреслення компонентів хвильової форми, таких як P-хвилі, комплекси QRS та T-хвилі. Цей метод продемонстрував високу чутливість (94,6 %) і специфічність (96,2 %) в автоматизованих завданнях сегментації та класифікації ЕКС, підвищуючи надійність виявлення аритмії в порівнянні з однонаправленими моделями (чутливість 89,3 %) (Yildirim, 2024).

Gated Recurrent Units (GRU), спрощений варіант LSTM, також забезпечують конкурентоспроможну продуктивність із зменшеною обчислювальною складністю приблизно на 30 %. Ці моделі успішно фіксують тонкі часові зміни, необхідні для діагностики аномалій інтервалу QT та інших серцевих дисфункцій, досягаючи точності 96,4 % і дозволяючи розгортання на мобільних пристроях завдяки своїй обчислювальній ефективності (Nurmaini, 2021).

Гібридні та вдосконалені моделі DL

Останні дослідження пропонують гібридні архітектури DL, що поєднують шари CNN та RNN, щоб скористатися перевагами вилучення просторових ознак CNN та моделювання часових послідовностей RNN. Такі інтегративні рамки дозволяють навчатися як на основі локальних морфологічних ознак, так і на основі довгострокових залежностей у даних ЕКС, що призводить до поліпшення надійності класифікації з точністю, що досягає 99,4 % на багатокласових наборах даних про аритмію (Janbhasha, 2023).

Для підвищення стійкості моделі було запроваджено об'єднання декількох мереж DL з порівнянними шарами, що дозволяє зменшити проблеми, пов'язані з тривалим часом навчання та залежністю від ручного вибору ознак. Оцінка багатокласових наборів даних про аритмію показує, що ці гібридні підходи перевершують традиційні методи за чутливістю (99,1 %), специфічністю (99,5 %) та обчислювальною ефективністю при належному оптимізуванні (Janbhasha, 2023).

У роботі Duong, L.T. (2023) досліджені альтернативні передові архітектури, такі як графічні нейронні мережі (GNN), що використовують графічні структури представлення ЕКС. GNN виявляють себе у вилученні реляційної інформації з даних ЕКС з декількома відведеннями і показали багатообіцяючі результати, перевершивши базові моделі на 12 % у завданнях локалізації інфаркту міокарда, потенційно надаючи більш інтерпретовані та клінічно значущі прогнози.

Підходи до ансамблевого та гібридного машинного навчання

Ансамблеві методи класифікації ЕКС

Ансамблеве навчання поєднує декілька окремих класифікаторів для досягнення вищої загальної ефективності. Такі методи, як усереднення, багінг та стекінг, використовують різноманітність та взаємодоповнювальні переваги складових моделей, підвищуючи надійність, особливо у випадку незбалансованих наборів даних ЕКС, де рідкісні класи аритмії вимагають ретельної ідентифікації.

Схеми голосування більшістю аґрегують прогнози для зменшення дисперсії та упередженості, забезпечуючи більш стабільні остаточні рішення. Наприклад, було показано, що поєднання AdaBoost з логістичною регресією та машинами опорних векторів дає ансамблеві класифікатори з підвищеною точністю (98,7 %) та F1-оцінками (0,962), що перевершують окремі моделі на публічних наборах даних ЕКС (Yildirim, 2024). Ці ансамблі є цінними в клінічних контекстах, де помилкові негативні результати можуть мати серйозні наслідки, при цьому частота помилкових негативних результатів зменшується на 18 % порівняно з окремими класифікаторами (Rath, 2022).

Поєднання експертних знань та ML

Інтеграція експертних клінічних знань з ML підвищує точність класифікації та інтерпретованість. Вибір експертних ознак на основі правил, заснований на фізіологічному розумінні, доповнює виявлення ознак ML на основі даних, забезпечуючи баланс між клінічною релевантністю та прогностичною здатністю.

Приклади порівняння моделей, навчених на основі ознак, обраних експертами, та ознак, обраних машиною, демонструють поліпшення на 15−22 % при інтеграції обох підходів, особливо для виявлення рідкісних захворювань. Такі гібридні моделі використовують переваги людської експертизи в даній галузі, що допомагає зменшити вплив шумних або нерелевантних ознак, тоді як алгоритми ML фіксують складні, неочевидні закономірності (Koubaa, 2024). Синергія між людським та машинним інтелектом сприяє клінічному застосуванню, надаючи моделі, які тісно відповідають встановленим діагностичним критеріям, одночасно використовуючи досягнення в галузі обчислювальної техніки (Reethunandh, 2023).

Гібридні моделі: інтеграція ML і DL

Гібридні фреймворки, що інтегрують класичний ML з підходами глибинного навчання, поєднують переваги обох. Традиційний ML може бути використаний для надійного вибору ознак або зменшення розмірності, щоб подавати компактні представлення в класифікатори DL, оптимізуючи використання ресурсів і зберігаючи точність вище 97 %.

Такі архітектури виявляються ефективними в системах моніторингу в реальному часі та носимих пристроях, які вимагають ефективних обчислень без втрати точності. Реалізації демонструють, що гібридні моделі пропонують вигідний компроміс між показниками продуктивності та обчислювальною складністю, при цьому час інференції скорочується на 40−60 % порівняно з чистими підходами DL, зберігаючи при цьому порівнянну точність (Śmigiel, 2024). Апаратна реалізація та валідаційні дослідження підкреслюють доцільність таких рішень, що дозволяють здійснювати безперервний моніторинг серцевої діяльності з енергоспоживанням нижче 10 мВт (Wasimuddin, 2020).

Загальні показники оцінки в класифікації ЕКС ML

Оцінка ефективності моделей класифікації ЕКС зазвичай включає такі показники, як точність, прецизійність, відтворюваність (чутливість), специфічність, F1-показник та площа під кривою ROC (AUC). Ці показники надають багатогранний погляд на поведінку класифікатора, враховуючи рівень справжніх позитивних результатів та вартість помилкових виявлень (Yildirim, 2024). Матриця плутанини є особливо інформативною в багатокласовій класифікації, висвітлюючи сильні та слабкі сторони моделі в різних класах і полегшуючи розуміння тенденцій помилкової класифікації, особливо в незбалансованих наборах даних, де сама по собі точність може вводити в оману (Philip, 2023).

Врахування незбалансованих даних є надзвичайно важливим при виборі метрик, що сприяє використанню збалансованих показників, таких як F1, для уникнення завищених оцінок ефективності. При клінічному застосуванні чутливість часто має пріоритет над специфічністю для станів, що загрожують життю, причому цільові порогові значення зазвичай встановлюються вище 95 % для таких станів, як фібриляція шлуночків (Hassaballah, 2023).

Обчислювальні обмеження становлять значну проблему для моніторингу ЕКС у реальному часі в носимих пристроях та пристроях IoT. Простіші класифікатори, такі як дерева рішень та випадкові ліси, забезпечують швидший час виведення висновків (2−5 мс), тоді як глибинні нейронні мережі вимагають більше обчислювальних ресурсів, але забезпечують вищу точність із часом виведення висновків 15−35 мс (Śmigiel, 2024).

Апаратні реалізації архітектур CNN підтвердили свою практичну здійсненність у портативних пристроях, коли складність мережі відповідним чином зменшується за допомогою квантування та обрізки. Компроміси між складністю та точністю є важливими для оптимізації швидкості реагування системи та енергоспоживання, при цьому квантовані моделі досягають 97 % від початкової точності, зменшуючи розмір моделі на 75 % (Wasimuddin, 2021). Підходи, що використовують двонаправлені мережі LSTM, також забезпечують баланс між продуктивністю та обчислювальним слідом, що робить їх привабливими для аналізу ЕКС на пристрої з часом висновків менше 20 мс (Nurmaini, 2021).

Обґрунтування актуальності дослідження

Незважаючи на високу точність глибинних нейронних мереж (96−99 % згідно з літературою), існують певні обмеження їх застосування в клінічній практиці, потреба у великих навчальних вибірках, висока обчислювальна складність, низька інтерпретованість рішень, неприйнятна для медичних застосувань. Тому актуальність даної роботи полягає в тому, що досліджується вплив попереднього опрацювання (StandardScaler + PCA) на ефективність алгоритмів виявлення аномалій для класифікації ЕКС; запропонований підхід дозволяє працювати з обмеженими навчальними вибірками (лише нормальні дані або один тип патології), що відповідає реаліям клінічної практики; інтегровано з математичним апаратом ЦВП для одночасного аналізу морфологічних та ритмічних ознак, чого не забезпечують існуючі методи.

Формулювання цілі статті

Для забезпечення всебічного дослідження ЕКС, що охоплює аналіз як структурних характеристик форми, так і часових параметрів серцевого ритму, було створено інформаційну технологію (ІТ) (Sverstiuk, 2025). В основі ІТ лежить математичний апарат циклічних випадкових процесів (ЦВП) (Lytvynenko, 2017). Ціллю даної роботи є експериментальне дослідження впливу методів попереднього опрацювання даних на ефективність алгоритмів виявлення аномалій при класифікації ЕКС за морфологічними та ритмічними ознаками, отриманими на основі математичного апарату ЦВП, та перевірка роботи блоку класифікації ритмічних ознак (БКРО) та блоку класифікації морфологічних ознак (БКМО) з використанням методів машинного навчання (нейронні мережі, SVM).

Обґрунтування вибору методів класифікації

У контексті розробленої ІТ на основі математичного апарату ЦВП, для блоків класифікації БКРО та БКМО було обрано алгоритми виявлення аномалій (anomaly detection) замість класичних підходів багатокласової класифікації або глибинних нейронних мереж. Таке рішення обґрунтовано наступними міркуваннями:

1) у задачі клінічної класифікації ЕКС часто виникає ситуація, коли навчальні дані доступні лише для нормального стану або одного конкретного типу патології, тоді як різноманітність патологічних станів є надзвичайно широкою. Алгоритми виявлення аномалій природно вирішують таку задачу “один клас проти решти” (one-class classification), навчаючись моделювати характеристики нормального стану без необхідності мати вичерпну вибірку всіх можливих патологій;

2) глибинні нейронні мережі (CNN, LSTM, bi-LSTM), згідно з аналізом літератури, демонструють високу точність (96−99 %) у задачах класифікації ЕКС, проте вони вимагають великих обсягів збалансованих навчальних даних (тисячі прикладів для кожного класу), мають високу обчислювальну складність та енергоспоживання;

3) обрані методи (OneClassSVM, IsolationForest, LOF, EllipticEnvelope) забезпечують високу інтерпретованість рішень, що критично важливо для медичних застосувань; можливість навчання на обмежених вибірках; швидкий час виведення висновків (2−5 мс), що дозволяє реалізувати моніторинг в режимі реального часу; природну інтеграцію з математичним апаратом ЦВП, де аналіз розкидів амплітудних значень та тривалостей сегментів безпосередньо відповідає логіці виявлення аномалій.

Дане дослідження є першим етапом верифікації запропонованої ІТ, де перевіряється принципова можливість класифікації за ритмічними та морфологічними ознаками. На наступному етапі планується проведення розширеного порівняльного експерименту з включенням архітектур CNN, LSTM та гібридних моделей на тих самих даних для об'єктивної оцінки переваг кожного підходу в контексті розробленої ІТ.

Виклад основного матеріалу

З метою оцінки практичної результативності запропонованої ІТ (рис. 1) виконано експериментальні випробування з використанням електрокардіографічних записів хворих, у яких виявлено аритмічні стани.

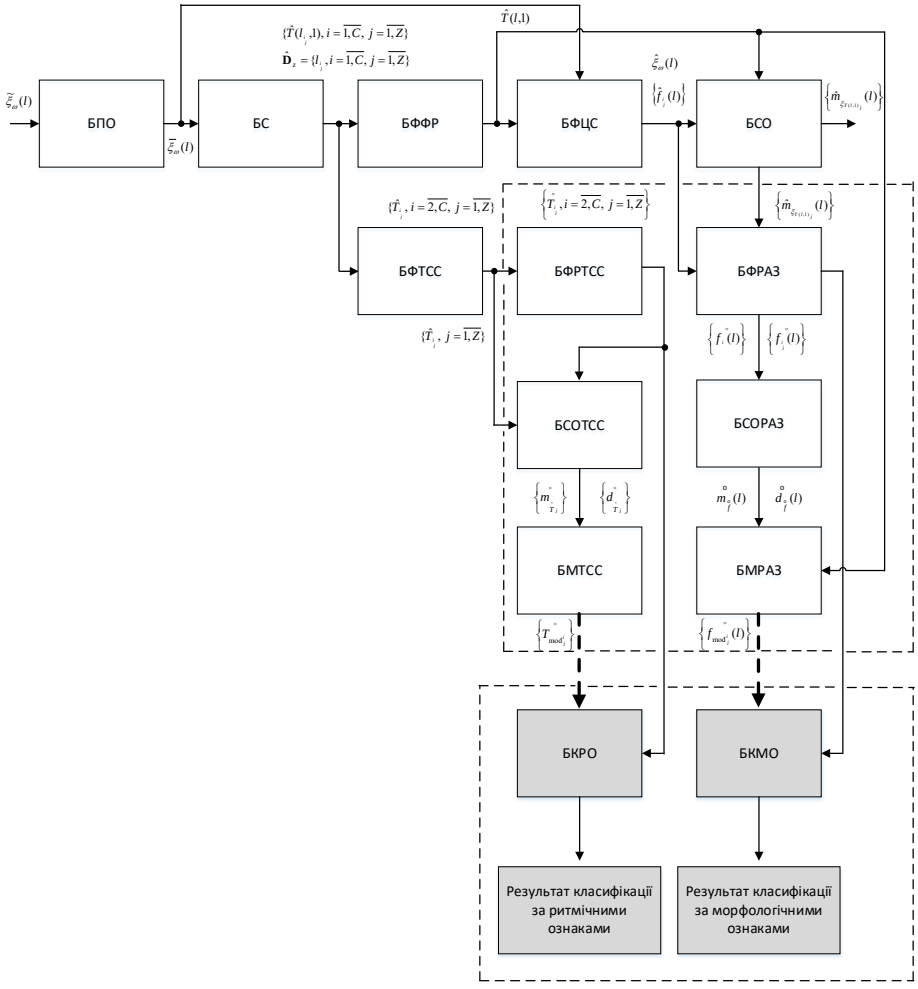


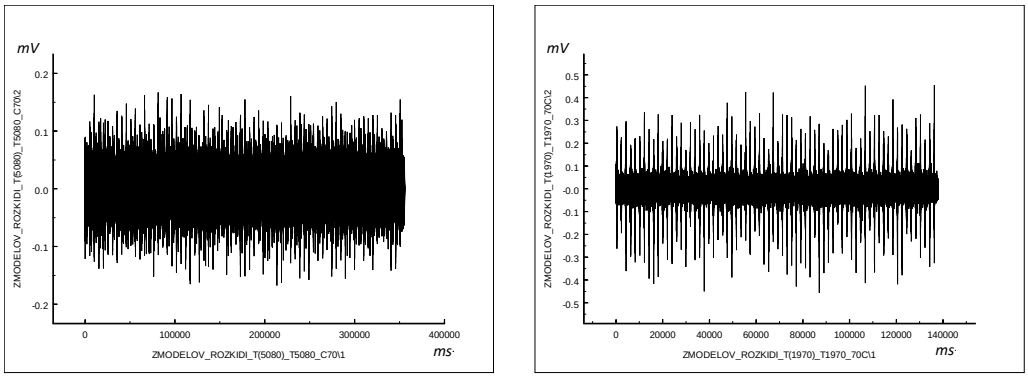
Рис. 1. Структурна схема ІТ аналізу ЕКС (Sverstiuk, 2025).

Для дослідження відібрано два типи передсердних аритмій з чітко вираженими змінами біоелектричної активності міокарда: 1) фібриляція передсердь (ФП) – проявляється неорганізованою біоелектричною активністю передсердного міокарда, зникненням зубців Р на електрокардіограмі та варіабельною тривалістю міжшлуночкових комплексів; 2) тріпотіння передсердь (ТП) – характеризується ритмічною високочастотною активацією передсердь з появою специфічних F-хвиль пилкоподібної конфігурації.

Розроблена ІТ складається з наступних функціональних блоків:

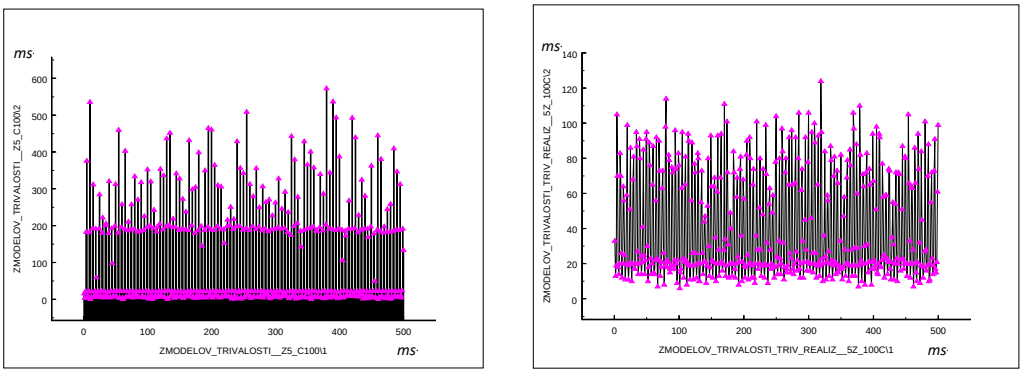
- БПО – блок попереднього опрацювання (здійснює фільтрацію та нормалізацію вхідного сигналу);
- БС – блок сегментації (виділяє окремі цикли ЕКС);
- БФФР – блок формування функції ритму (визначає часові параметри серцевого ритму);
- БФЦС – блок формування циклічного сигналу (перетворює ЕКС у циклічну структуру);
- БСО – блок статистичного опрацювання (обчислює статистичні характеристики);
- БФТСС – блок формування тривалостей сегментної структури (визначає тривалості сегментів ЕКС);
- БФРТСС – блок формування розкидів тривалостей сегментної структури (аналізує варіабельність часових інтервалів);
- БФРАЗ – блок формування розкидів амплітудних значень (визначає варіабельність амплітуд);
- БСОТСС – блок статистичного опрацювання тривалостей сегментної структури (статистичний аналіз часових параметрів);
- БСОРАЗ – блок статистичного опрацювання розкидів амплітудних значень (статистичний аналіз амплітудних параметрів);
- БМТСС – блок моделювання тривалостей сегментної структури (генерує модельні часові характеристики);
- БМРАЗ – блок моделювання розкидів амплітудних значень (генерує модельні амплітудні характеристики);
- БКРО – блок класифікації ритмічних ознак (класифікує патології за ритмічними порушеннями);
- БКМО – блок класифікації морфологічних ознак (класифікує патології за морфологічними змінами).

Для навчання класифікаторів використовувались дані, подані на рис. 2−3. Дані для навчання були згенеровані на основі статистичних характеристик математичного сподівання та дисперсії з діагнозами фібриляції та тріпотіння передсердь (порушення морфологічного характеру та порушення ритму). Валідація змодельованих даних проводилась шляхом порівняння спектральних характеристик з реальними записами (коефіцієнт кореляції >0,85) та експертної оцінки кардіологом на предмет збереження діагностичних ознак.



а) *б)*

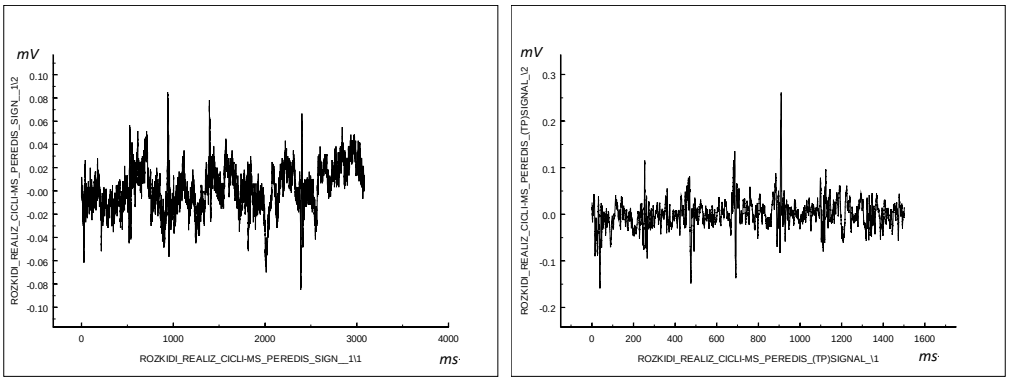
Рис. 2. Дані для навчання класифікаторів (патології що проявляються у порушенні морфології):а) ФП; б) ТП



а) *б)*

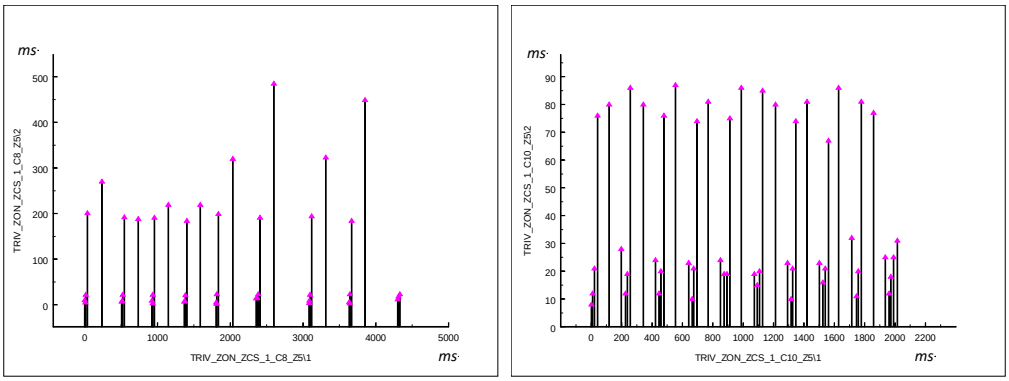
Рис. 3. Дані для навчання класифікаторів
(патології що проявляються у порушенні ритмі):а) ФП; б) ТП

Тестові реалізації розкидів амплітудних значень та тривалостей сегментної структури ЕКС для верифікації класифікаторів представлені на рис. 4−5.



а) *б)*

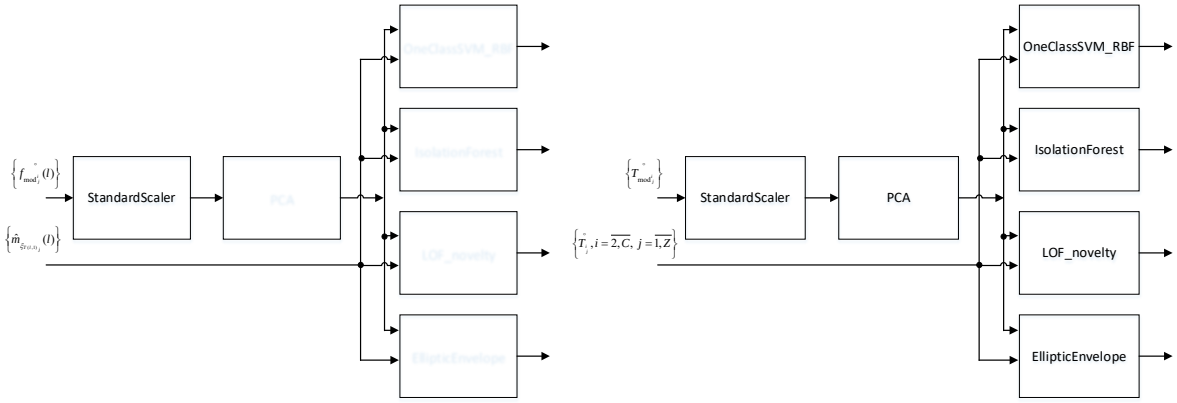
Рис. 4. Розкиди амплітудних значень
(для перевірки класифікаторів:а) ФП; б) ТП



а) *б)*

Рис. 5. Тривалості сегментної структури ЕКС
(для перевірки класифікаторів): а) ФП; б) ТП

Для класифікації патологій використовували, як попереднє опрацювання вхідних даних, так і навчались класифікатори без застосування попереднього опрацювання, зокрема, такі структурні схеми зображені на рис. 6.



а) *б)*

Рис. 6. Структурна схема опрацювання сигналу:а) при патології, що проявляється у порушенні морфології; б) при патології, що проявляється у порушенні ритму

Стандартизація набору даних є загальною вимогою для багатьох оцінювачів машинного навчання. Вони можуть поводитися погано, якщо окремі ознаки не виглядають більш-менш як стандартні нормально розподілені дані (наприклад, гаусівські з 0 середнім значенням та одиничною дисперсією).

Наприклад, багато елементів, що використовуються в цільовій функції алгоритму навчання (такі як ядро RBF методів опорних векторів або регуляризатори L1 та L2 лінійних моделей), припускають, що всі ознаки зосереджені навколо 0 та мають дисперсію одного порядку. Якщо ознака має дисперсію, яка на порядок більша за інші, вона може домінувати в цільовій функції та зробити оцінювач нездатним правильно навчатися на основі інших ознак, як очікувалося. Тому з цією метою ми використовували StandardScaler.

Ми також застосовували Principal component analysis (PCA) – метод зниження розмірності даних, який використовується для спрощення аналізу та візуалізації, зберігаючи при цьому найбільш важливу інформацію. PCA перетворює вихідні змінні в новий набір незалежних компонентів (головних компонент), які відсортовані за важливістю (дисперсією), не корелюють між собою, пояснюють максимальну варіативність у даних.

Серед вибраних для дослідження класифікаторів слід назвати наступні:

OneClassSVM (RBF) – це алгоритм виявлення аномалій (outlier detection), який використовує метод опорних векторів (SVM) з радіально-базисною функцією (RBF kernel). Модель навчається тільки на “нормальних” даних (один клас). Вона будує межу (гіперплощину) навколо нормальних точок, а нові точки, що потрапляють за межу, вважаються аномаліями.

Isolation Forest – це алгоритм виявлення аномалій, який базується на ідеї ізоляції аномальних точок замість моделювання нормальних даних. Аномальні точки легше ізолювати (відокремити) від нормальних, бо вони рідкісні та сильно відрізняються.

Алгоритм будує випадкові дерева (random trees), які розбивають дані:

- чим коротший шлях до ізоляції точки – тим більше ймовірність, що це аномалія;
- нормальні точки зазвичай потребують більшої кількості розділень.

LOF (Local Outlier Factor) – це алгоритм для виявлення аномалій, який порівнює густину даних у локальному оточенні кожної точки. Для кожної точки обчислюється густина сусідів (k-nearest neighbors). Аномальні точки зазвичай мають нижчу локальну густину, ніж сусіди.

Цей метод має переваги:

- виявляє локальні аномалії, які інші алгоритми можуть пропустити;
- добре працює з нелінійними даними;
- природно підходить для задач, де важлива густина.

EllipticEnvelope – це алгоритм для виявлення аномалій, який базується на статистичному підході. Він припускає, що нормальні дані мають багатовимірний нормальний розподіл (Gaussian distribution) і будує еліптичну межу, яка охоплює більшість нормальних точок.

Обчислює середнє значення (mean) та коваріаційну матрицю (covariance matrix) для даних. Формує еліпсоїд у багатовимірному просторі, який описує "нормальні" точки. Точки, що потрапляють за межі еліпсоїда, вважаються аномаліями. Простий і швидкий для даних, які приблизно мають нормальний розподіл. Добре працює для багатовимірних даних, якщо припущення про Gaussian розподіл справджується.

Отримані результати подані в табл. 1−4. Зокрема в табл. 1 застосування класифікаторів без попереднього опрацювання вхідних даних.

Таблиця 1

Фрагменти отриманих результатів застосування класифікаторів (без попереднього опрацювання) для патології, що проявляється у порушенні морфології (ФП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	Попередня оцінка (n=6)
OneClassSVM_RBF	1	−1	1	1	−1	−1	50 %
IsolationForest	1	−1	1	1	−1	−1	50 %
LOF_novelty	1	1	1	1	1	−1	83 %
EllipticEnvelope	–	–	–	–	–	–	–

де: «1» – класифікатор коректно визначив приналежність ЕКС до навченого класу, «−1» – некоректно визначив, «–» – не визначив взагалі.

Таблиця 2

Фрагменти отриманих результатів застосування класифікаторів (з попереднім опрацюванням) для патології, що проявляється у порушенні морфології (ФП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	Попередня оцінка (n=6)
OneClassSVM_RBF	1	1	1	1	1	1	100 %
IsolationForest	1	1	1	1	1	1	100 %
LOF_novelty	1	1	1	1	1	1	100 %
EllipticEnvelope	1	1	1	1	1	1	100 %

Як бачимо з отриманих результатів застосування попереднього опрацювання дає кращі результати класифікації.

Таблиця 3

Фрагменти отриманих результатів застосування класифікаторів (без попереднього опрацювання) для патології, що проявляється у порушенні морфології (ТП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	ЕКС 7	Попередня оцінка (n=7)
OneClassSVM_RBF	−1	1	1	1	−1	1	1	71 %
IsolationForest	−1	1	1	1	1	−1	1	71 %
LOF_novelty	1	1	1	1	−1	1	1	86 %
EllipticEnvelope	1	1	1	1	1	1	1	100 %

Таблиця 4

Фрагменти отриманих результатів застосування класифікаторів (з попереднім опрацюванням) для патології, що проявляється у порушенні морфології (ТП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	ЕКС 7	Попередня оцінка (n=7)
OneClassSVM_RBF	1	1	1	1	1	1	1	100 %
IsolationForest	1	1	1	1	1	1	1	100 %
LOF_novelty	1	1	1	1	1	1	1	100 %
EllipticEnvelope	−1	−1	−1	−1	1	1	−1	28 %

Якщо порівнювати отримані результати, то слід відмітити, що серед обраних класифікаторів найкраще справляється LOF_novelty. Саме він в більшості випадків при різних патологіях дає кращі результати і оцінка роботи класифікатора в межах від 83 до 100 %. Саме тому його можна застосовувати в аналізі розкидів амплітудних значень для класифікації патологій. У перспективі необхідно буде провести експерименти з аналізом інших видів патологій, які проявляються у порушенні морфології.

Тепер розглянемо класифікацію патологій, які проявляються у порушенні ритму. Отримані результати подані в табл. 5−8. Зокрема в таблиці 5 застосування класифікаторів без попереднього опрацювання вхідних даних для випадку патології, що проявляється у порушенні ритму.

Таблиця 5

Фрагменти отриманих результатів застосування класифікаторів (без попереднього опрацювання) для патології, що проявляється у порушенні ритму (ФП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	Попередня оцінка (n=6)
OneClassSVM_RBF	1	1	−1	−1	1	1	71 %
IsolationForest	1	1	1	−1	1	1	86 %
LOF_novelty	1	1	1	−1	1	1	86 %
EllipticEnvelope	1	1	1	−1	1	1	86 %

Таблиця 6

Фрагменти отриманих результатів застосування класифікаторів (з попереднім опрацюванням) для патології, що проявляється у порушенні ритму (ФП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	Попередня оцінка (n=6)
OneClassSVM_RBF	1	1	1	−1	1	1	86 %
IsolationForest	1	1	1	−1	1	1	86 %
LOF_novelty	1	1	1	−1	1	1	86 %
EllipticEnvelope	1	1	1	−1	1	1	86 %

У випадку визначення приналежності циклу до відповідного класу (ЕКС 4) застосовані методи класифікації не вирішили поставлену задачу, це пов’язано перш за все з аномальністю сигналу (наявністю артефактів).

Таблиця 7

Фрагменти отриманих результатів застосування класифікаторів (без попереднього опрацювання) для патології, що проявляється у порушенні ритму (ТП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	Попередня оцінка (n=6)
OneClassSVM_RBF	1	1	1	1	1	−1	89 %
IsolationForest	1	1	1	1	1	−1	89 %
LOF_novelty	1	1	1	1	1	−1	89 %
EllipticEnvelope	1	1	1	1	1	−1	89 %

Таблиця 8

Фрагменти отриманих результатів застосування класифікаторів (з попереднім опрацюванням) для патології, що проявляється у порушенні ритму (ТП)

Класифікатори	ЕКС 1	ЕКС 2	ЕКС 3	ЕКС 4	ЕКС 5	ЕКС 6	Попередня оцінка (n=6)
OneClassSVM_RBF	−1	1	1	1	1	−1	78 %
IsolationForest	1	1	1	1	1	−1	89 %
LOF_novelty	1	1	1	1	1	−1	89 %
EllipticEnvelope	1	1	1	1	1	−1	89 %

У випадку класифікації для сигналів з патологією, що проявляється у порушенні ритму, кращими себе показали IsolationForest, LOF_novelty, EllipticEnvelope з однаковою ефективністю на рівні 89 %. У випадку опрацювання ЕКС 6 також всі застосовані методи не справились із завданням. У подальших дослідженнях необхідно застосувати даний підхід до аналізу сигналів, які містять інші типи патологій, що проявляються у порушенні ритму, з метою визначення, який із обраних класифікаторів буде коректно справлятись з поставленим завданням у порівнянні з іншими.

Висновки

У роботі проведено експериментальне дослідження ефективності методів машинного навчання для класифікації ЕКС за ритмічними та морфологічними ознаками в рамках розробленої ІТ на основі математичного апарату ЦВП. Встановлено критичну важливість попереднього опрацювання даних для класифікації морфологічних порушень. Застосування комбінації StandardScaler та PCA підвищило точність класифікації фібриляції передсердь з 50−83 % до 100 % для всіх досліджених класифікаторів. Це підтверджує гіпотезу про необхідність нормалізації та зниження розмірності даних при аналізі амплітудних характеристик ЕКС. Слід зазначити, що представлені результати отримані на обмеженій тестовій вибірці (6−7 ЕКС для кожного типу патології), що не дозволяє робити остаточні висновки про статистичну достовірність. Наведені відсотки точності (50−100 %) слід розглядати як попередні оцінки, що демонструють принципову працездатність підходу.

Виявлено, що алгоритм LOF (Local Outlier Factor) з налаштуванням novelty демонструє найбільш стабільні результати для різних типів патологій. При класифікації морфологічних порушень без попереднього опрацювання LOF досягав точності 83−86 %, що перевищує показники інших методів. Це свідчить про ефективність підходу на основі локальної густини для виявлення аномалій в ЕКС.

Для класифікації ритмічних порушень методи IsolationForest, LOF_novelty та EllipticEnvelope показали однаково високу ефективність (89 %) при аналізі тріпотіння передсердь. Водночас, попереднє опрацювання даних не призвела до суттєвого покращення результатів, а в деяких випадках навіть погіршила їх, що вказує на специфіку часових характеристик ЕКС, які можуть втрачати інформативність при трансформаціях.

Порівняння отриманих результатів з даними публікацій щодо застосування глибинних нейронних мереж показує конкурентоспроможність запропонованого підходу. Згідно з оглядом літератури, архітектури CNN досягають точності 98,5 % на наборі даних MIT-BIH для класифікації аритмій, двонаправлені LSTM забезпечують чутливість 94,6 % та специфічність 96,2 %, а гібридні CNN-RNN моделі демонструють точність до 99,4 % на багатокласових наборах даних. Отримані в нашому дослідженні результати (89−100 % залежно від типу патології та методу) знаходяться в тому ж діапазоні ефективності, але досягаються при суттєво менших обчислювальних витратах та кращій інтерпретованості результатів. Водночас, для остаточних висновків щодо переваг того чи іншого підходу необхідне проведення прямого порівняльного експерименту на ідентичних наборах даних з урахуванням специфіки задачі класифікації в контексті розробленої ІТ на основі ЦВП.

Перспективи подальших досліджень включають:

- проведення прямого порівняльного експерименту з архітектурами CNN, LSTM/bi-LSTM та гібридними CNN-RNN моделями на тих самих наборах даних ЕКС для об’єктивної оцінки переваг кожного підходу з точки зору точності, обчислювальної складності та інтерпретованості;
- розширення спектру аналізованих патологій, включаючи шлуночкові аритмії, блокади провідності та ішемічні зміни;
- дослідження гібридних архітектур, що поєднують методи виявлення аномалій з елементами глибинного навчання для автоматичного вилучення ознак;
- розроблення адаптивних алгоритмів автоматичного вибору методів опрацювання та класифікації залежно від типу порушення та характеристик сигналу;
- для статистично достовірної оцінки необхідне тестування на вибірці не менше 100−200 ЕКС для кожного класу патології з розрахунком довірчих інтервалів та проведенням кросвалідації.
- дослідження можливості інтеграції запропонованого підходу з носимими медичними пристроями для безперервного моніторингу серцевої діяльності в реальному часі.

СПИСОК ЛІТЕРАТУРИ

Abdul Razak, S. F., Sayed Ismail, S. N. M., Yogarayan, S., Azli Abdullah, M. F., Kamis, N. H., & Abdul Aziz, A. (2023). Comparative Study of Machine Learning Algorithms in Classifying HRV for the Driver's Physiological Condition. Civil Engineering Journal, 9(9), 2272–2285. https://doi.org/10.28991/cej-2023-09-09-013

Agrawal, R. K., Sewani, R. R., Delen, D., & Benjamin, B. (2022). A machine learning approach for classifying healthy and infarcted patients using heart rate variabilities derived vector magnitude. Healthcare Analytics, 2, 100121. https://doi.org/10.1016/j.health.2022.100121

Coronado-Reyes, O. I., Téllez-Anguiano, A. C., Castro-Pimentel, L. A., & Gutierrez-Gnecchi, J. A. (2025). Non-invasive Hyperglycemia Detection via Electrocardiogram Using Discrete Wavelet Transform and Machine Learning. Cureus. https://doi.org/10.7759/cureus.80548

Duong, L. T., Doan, T. T. H., Chu, C. Q., & Nguyen, P. T. (2023). Fusion of edge detection and graph neural networks to classifying electrocardiogram signals. Expert Systems with Applications, 225, 120107. https://doi.org/10.-1016/j.eswa.2023.120107

Gupta, T. R., & Nandhini, D. U. (2023). Learning Techniques of ECG Arrhythmia Classification: A Review. In 2023 1st International Conference on Cognitive Computing and Engineering Education (ICCCEE) (pp. 1–5). IEEE. https://doi.org/10.1109/icccee55951.2023.10424554

Hassaballah, M., Wazery, Y. M., Ibrahim, I. E., & Farag, A. (2023). ECG Heartbeat Classification Using Machine Learning and Metaheuristic Optimization for Smart Healthcare Systems. Bioengineering, 10(4), 429. https://doi.org/10.3390/bioengineering10040429

Janbhasha, S., & Bhavanam, S. N. (2023). A Comparative Analysis of the Feature Selection Process Using Deep Learning Methods for Arrhythmia. In 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) (Vol. 119, pp. 1–6). IEEE. https://doi.org/10.-1109/sceecs57921.2023.10063070

Koubaa, H., Boujelben, M., Hdidane, M., & Kallel, R. (2024). An ECG Data Feature Selection Comparison: Machine Learning and Expert Doctors. In 2024 IEEE/ACS 21st International Conference on Computer Systems and Applications (AICCSA) (pp. 1–5). IEEE. https://doi.org/10.1109/aiccsa63423.2024.10912630

Kumar, S., & Veer, K. (2023). Evaluation of Current Trends in Biomedical Applications Using Soft Computing. Current Bioinformatics, 18(9), 693–714. https://doi.org/10.2174/1574893618666230706112826

Kumar M., A., & Chakrapani, A. (2022). Classification of ECG signal using FFT based improved Alexnet classifier. PLOS ONE, 17(9), e0274225. https://doi.org/10.1371/journal.pone.0274225

Lytvynenko Ya.V. (2017). The method of segmentation of stochastic cyclic signals for the problems of their processing and modeling. Journal of Hydrocarbon Power Engineering, Oil and Gas Measurement and Testing. 4(2), 93–103.

Nurmaini, S., Tondas, A. E., Darmawahyuni, A., Rachmatullah, M. N., Effendi, J., Firdaus, F., & Tutuko, B. (2021). Electrocardiogram signal classification for automated delineation using bidirectional long short-term memory. Informatics in Medicine Unlocked, 22, 100507. https://doi.org/10.1016/j.imu.2020.100507

Philip, A. M., & Hemalatha, S. (2023). A Performance Analysis Of Advanced Machine Learning Techniques For Arrhythmia Detection By Categorizing Ecg Signals. In 2023 International Conference for Technological Engineering and its Applications in Sustainable Development (ICTEASD) (pp. 7–12). IEEE. https://doi.org/10.1109/icteasd57136.2023.10585040

Rath, A., Mishra, D., & Panda, G. (2022). Imbalanced ECG signal-based heart disease classification using ensemble machine learning technique. Frontiers in Big Data, 5. https://doi.org/10.3389/fdata.2022.1021518

Reethunandh, H R, Y., Venkata Sandeep, T. S., K R, S. N., & Chandrashekar, H. M. (2023). Classification of ECG Arrhythmia Using a Convolution Neural Network. In 2023 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES) (pp. 1–8). IEEE. https://doi.org/10.1109/icsses 58299.2023.10200579

Śmigiel, S., Topoliński, T., Ledziński, D., & Andrysiak, T. (2024). The ECG Signal Monitoring System Using Machine Learning Methods and LoRa Technology. Pomiary Automatyka Robotyka, 28(2), 21–36. https://doi.org/10.14313/par_252/21

Sverstiuk A., & Mosiy L. (2025). Information technology for electrocardiographic signal analysis based on mathematical models of temporal and amplitude variability. Computer Systems and Information Technologies, (2), 36–44. https://doi.org/10.31891/csit-2025-2-4

Sverstiuk A., & Mosiy, L. (2025). Інформаційна технологія опрацювання та аналізу електрокардіосигналів з врахуванням їх морфологічних та ритмічних ознак. Computer-integrated technologies: education, science, production, (60), 40-52. https://doi.org/10.36910/6775-2524-0560-2025-60-04

Ullah, A., & Khan, M. S. (2024). Classification of Sleep Stages Using Single-Channel ECG Signals: A Comparative Analysis of Machine Learning and Deep Learning Methods. In 2024 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC) (pp. 1–6). IEEE. https://doi.org/10.-1109/icspcc62635.2024.10770354

Wasimuddin, M., Elleithy, K., Abuzneid, A., Faezipour, M., & Abuzaghleh, O. (2020). Stages-Based ECG Signal Analysis From Traditional Signal Processing to Machine Learning Approaches: A Survey. IEEE Access, 8, 177782–177803. https://doi.org/10.1109/access.2020.3026968

Wasimuddin, M., Elleithy, K., Abuzneid, A., Faezipour, M., & Abuzaghleh, O. (2021). Multiclass ECG Signal Analysis Using Global Average-Based 2-D Convolutional Neural Network Modeling. Electronics, 10(2), 170. https://doi.org/10.3390/electronics10020170

Yildirim, B. E., Taskiran, M., & Nur Bekiroglu, K. (2024). Ensemble Learning Methodologies for Electrocardiogram Analysis: A Comparative Study. In 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) (pp. 1–6). IEEE. https://doi.org/10.1109/iitcee59897.2024.10467494

Yue, Y., Chen, C., Liu, P., Xing, Y., & Zhou, X. (2021). Automatic Detection of Short-Term Atrial Fibrillation Segments Based on Frequency Slice Wavelet Transform and Machine Learning Techniques. Sensors, 21(16), 5302. https://doi.org/10.3390/s21165302

MACHINE LEARNING METHODS FOR CLASSIFICATION OF ELECTROCARDIOGRAPHIC SIGNALS BASED ON RHYTHMIC AND MORPHOLOGICAL FEATURES

Lyubomyr Mosiy[1], Andriy Sverstiuk[1,2]

[1] Ternopil Ivan Puluj National Technical University,
Department of Computer Science, Ternopil, Ukraine
[2] Ternopil Ivan Horbachevsky National Medical University,
Department of Medical Informatics, Ternopil, Ukraine
E-mail: lmosiy@gmail.com, ORCID: 0009-0000-9778-331X
E-mail: sverstyuk@tdmu.edu.ua, ORCID: 0000-0001-8644-0776

© *Mosiy L., Sverstiuk A.*

The article presents an experimental study of the effectiveness of machine learning methods for classifying electrocardiographic signals by rhythmic and morphological features using information technology based on the mathematical apparatus of cyclic random processes. The problem of automated detection of atrial arrhythmias is considered, particularly atrial fibrillation and atrial flutter, which are characterized by complex changes in both ECG wave morphology and cardiac cycle time intervals.

Four anomaly detection algorithms were investigated for classifying pathological conditions: OneClassSVM with radial basis function, IsolationForest, Local Outlier Factor (LOF), and EllipticEnvelope. The impact of data preprocessing methods (StandardScaler for standardization and PCA for dimensionality reduction) on classification accuracy was analyzed. Experimental results showed that preprocessing is critically important for morphological disorder classification, increasing accuracy from 50-83 % to 100 % for atrial fibrillation. The LOF algorithm demonstrated the most stable results (83-100 %) for different types of pathologies. For rhythm disorder classification, the IsolationForest, LOF, and EllipticEnvelope methods showed equally high efficiency (89 %), while preprocessing did not lead to significant improvement in results.

Keywords - electrocardiographic signal modeling, information technology, cyclic discrete random process, amplitude-temporal characteristics, temporal rhythm function, amplitude variability, signal classification, artificial intelligence (AI), machine learning system (MLS).