

Міністерство освіти і науки України

Тернопільський національний технічний університет імені Івана Пулюя

(повне найменування вищого навчального закладу)

Факультет комп'ютерно-інформаційних систем і програмної інженерії

(назва факультету)

Кафедра систем штучного інтелекту та аналізу даних

(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

Магістр

(назва освітнього ступеня)

на тему: Технологія проектування архітектури трансформера

для класифікації багатовимірних часових рядів

Виконав: студент 6 курсу, групи САМ-61

спеціальності 124 Системний аналіз

(шифр і назва спеціальності)

(підпис) Лісовий В.Г.

(прізвище та ініціали)

Керівник доц. Яцишин В.В.

(підпис) (прізвище та ініціали)

Нормоконтроль ст. викл. Гладьо С.В.

(підпис) (прізвище та ініціали)

Завідувач доц. Яцишин В.В.

кафедри (підпис) (прізвище та ініціали)

Рецензент доц. Боднарчук І.О.

(підпис) (прізвище та ініціали)

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

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних систем штучного інтелекту та аналізу даних
(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

доц. Яцишин В.В.

(підпис)

(прізвище та ініціали)

« »

20__ р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

на здобуття освітнього ступеня магістр
(назва освітнього ступеня)

за спеціальністю 124 Системний аналіз
(шифр і назва спеціальності)

студенту Лісовому Володимиру Григоровичу

1. Тема роботи Технологія проектування архітектури трансформера
для класифікації багатовимірних часових рядів

Керівник роботи Яцишин Василь Володимирович, к.т.н, доц, зав. каф. СА
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом по університету від «07» листопада 2025 року № 4/7-951

2. Термін подання студентом роботи 15.12.2025

3. Вихідні дані до роботи наукові літературні джерела

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

1. Аналіз предметної області.

2. Теоретична частина.

3. Практична частина

4. Охорона праці та безпека в надзвичайних ситуаціях

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

1. Тема роботи. 2. Актуальність. 3. Мета, задачі, об'єкт, предмет дослідження

4. Наукова новизна, практичне значення, апробація роботи. 5. Основні завдання, пов'язані з аналізом часових рядів. 6. Стратегії токенизації. 7. Архітектура розроблюваної моделі

8. Вхідні дані, пошук закономірностей. 9. Набір даних UEA

10. Приклади датасетів із UEA. 11. Програмні засоби, використані для розробки.

12. Послідовність роботи інструменту автоматичного підбору гіперпараметрів

13. Інтерфейс інструменту. 14. Алгоритм тестування. 15, 16. Точність за всіма прикладами..

SOTA (DuFormer) та розробленої моделі. 17 Вплив особливостей розробленої моделі.

18. Основні результати дослідження.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Окіпний І.Б., зав. каф. МТ		
Безпека в надзвичайних ситуаціях	Стручок В.С., ст. викл. каф. ОХ		

7. Дата видачі завдання _____ 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1	Затвердження теми кваліфікаційної роботи	07.11.25	Виконано
2	Аналіз літературних джерел	08.11.25–13.11.25	Виконано
3	Обґрунтування актуальності дослідження	14.11.25–18.11.25	Виконано
4	Аналіз предмету дослідження та предметної області	19.11.25–22.11.25	Виконано
5	Проведення дослідження методів та засобів аналітичного опрацювання даних	23.11.25 –26.11.25	Виконано
6	Оформлення розділу «Аналіз предметної області»	27.11.25–29.11.25	Виконано
7	Оформлення розділу «Теоретична частина»	30.11.25–02.12.25	Виконано
8	Оформлення розділу «Практична частина»	03.12.25–06.12.25	Виконано
9	Оформлення розділу «Охорона праці та безпека в надзвичайних ситуаціях»	07.12.25–09.12.25	Виконано
10	Нормоконтроль	10.12.25–13.12.25	Виконано
11	Перевірка на плагіат	14.12.25–16.12.25	Виконано
12	Попередній захист роботи	15.12.25	Виконано
13	Захист кваліфікаційної роботи	23.12.25	

Студент

_____ (підпис)

Лісовий В.Г.

_____ (прізвище та ініціали)

Керівник роботи

_____ (підпис)

Яцишин В.В.

_____ (прізвище та ініціали)

АНОТАЦІЯ

Технологія проектування архітектури трансформера для класифікації багатовимірних часових рядів // Кваліфікаційна робота освітнього рівня «Магістр» // Лісовий Володимир Григорович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем та програмної інженерії, кафедра систем штучного інтелекту та аналізу даних, група САМ-61 // Тернопіль, 2025 // С. – 71, рис. – 44, табл.– 5, слайдів – 18, додат. – 1, бібліогр. – 53.

Ключові слова: багатовимірний часовий ряд, гіперпараметри, енкодер, патч, трансформер

Кваліфікаційна робота спрямована на створення архітектури трансформера для класифікації багатовимірних часових рядів, яка дає змогу враховувати взаємозв'язки між усіма вимірами, а також і вздовж кожного виміру часового ряду. Архітектура дозволить виявляти закономірності багатовимірних часових рядів та покращити результати класифікації для різних завдань.

У першій частині наведено особливості проведення аналізу часових рядів із використанням методів машинного навчання та трансформерів для їх класифікації. Описані основні підходи до вирішення задачі класифікації багатовимірних часових рядів. Також у цій частині виявляються основні недоліки існуючих методів.

У другій частині наведена архітектура розробленої моделі. Докладно описано дані з тестового набору для оцінки розробленої моделі.

У третій частині виявляються основні характеристики використовуваних даних, описується розроблений інструмент автоматичного підбору гіперпараметрів моделі. Докладно розкриваються кроки тестування та його результати. Розділ закінчується виявленням впливу особливостей розробленої архітектури на результати тестування.

ANNOTATION

Transformer architecture design technology for multidimensional time series classification // Lisovyi Volodymyr // Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Artificial Intelligence Systems and Data Analysis, SAm-61 group // Ternopil, 2025 // P. - 71, Fig. - 44, Table – 5, Slide - 18, References - 53.

Keywords: multidimensional time series, hyperparameters, encoder, patch, transformer

Thesis is aimed at creating a transformer architecture for classifying multidimensional time series, which allows taking into account the relationships between all dimensions, as well as along each dimension of the time series. The architecture will allow identifying patterns in multidimensional time series and improving classification results for various tasks.

The first part presents the features of time series analysis using machine learning methods and transformers for their classification. The main approaches to solving the problem of classifying multidimensional time series are described. This part also reveals the main shortcomings of existing methods.

The second part presents the architecture of the developed model. The data from the test set for evaluating the developed model is described in detail.

The third part reveals the main characteristics of the data used, describes the developed tool for automatic selection of model hyperparameters. The testing steps and its results are disclosed in detail. The section ends with identifying the influence of the features of the developed architecture on the testing results.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ

Feed-forward layer – повнозв'язковий шар (прямого поширення).

Patching – стратегія токенизації по локальних областях.

Self-attention – механізм самоуваги.

SOTA (state-of-the-art) – найсучасніший (передовий досвід).

TPT (Time Point Token) – стратегія токенизації за точками.

TST (time-series transformer) – трансформер часових рядів.

TVT (Time Variable Token) – стратегія токенизації за вимірами.

UEA (University of East Anglia) – Університет Східної Англії.

Аналіз часових рядів – сукупність математико-статистичних методів для виявлення структури даних, що змінюються в часі, і для прогнозування їх майбутніх значень.

БЧР – багатовимірні часові ряди.

Кластеризація часових рядів – групування часових рядів на основі їх подібності або загальних характеристик.

Прогнозування часових рядів – передбачення майбутніх значень часового ряду на основі його попередніх значень.

Трансформер – модель машинного навчання на основі механізму уваги.

ЗМІСТ

Вступ.....	9
1 Аналіз предметної області.....	11
1.1 Проведення аналізу часових рядів із використанням методів машинного навчання	11
1.2 Застосування моделей трансформерів для класифікації багатовимірних часових рядів.....	14
1.3 Стратегії токенизації.....	17
1.3.1 TPT.....	18
1.3.2 TVT	18
1.3.3 Patching.....	18
1.4 Недоліки застосування існуючих моделей трансформерів для класифікації багатовимірних часових рядів	19
1.5 Висновки до першого розділу.....	21
2 Теоретична частина.....	22
2.1 Архітектура розроблюваної моделі.....	22
2.2 Опис даних для оцінки розробленої моделі	24
2.3 Висновки до другого розділу	40
3 Практична частина	41
3.1 Інструмент автоматичного підбору гіперпараметрів	41
3.1.1 Алгоритм роботи інструменту.....	41
3.1.2 Інтерфейс інструменту	43
3.1.3 Алгоритм використання інструменту	46
3.2 Тестування розробленої моделі	47
3.3 Вплив особливостей розробленої моделі	50
3.3.1 Вплив розміру патчу.....	50
3.3.2 Вплив кількості енкодерів.....	53
3.4 Висновки до третього розділу	55
4 Охорона праці та безпека в надзвичайних ситуаціях	56

4.1 Закордонний досвід організації охорони праці в ІТ-компаніях	56
4.2 Оцінка дії електромагнітного імпульсу на елементи комп'ютерної системи.	60
4.3 Висновки до четвертого розділу.....	63
Висновки	64
Перелік джерел	66
Додатки	

ВСТУП

Актуальність теми. Класифікація БЧР є актуальним завданням у світі. До неї відносяться велика кількість розв'язуваних проблем у різних сферах діяльності. Для вирішення цього завдання вченими були винайдені різні методи та підходи, які дозволяють вирішити задачу класифікації БЧР.

Основна складність проведення аналізу БЧР у тому, що необхідно вловити закономірності як серед прикладів, а й усередині кожного прикладу між вимірами і між точками в часових рядах. З появою моделей трансформерів значно зросла якість аналізу послідовностей. Архітектура даних моделей дозволяє вловити закономірності, які раніше не вдалося виявити.

Мета дослідження: розробка архітектури трансформера для класифікації БЧР, що дозволяє враховувати взаємозв'язки між вимірами та вздовж кожного виміру часового ряду. Дана архітектура дозволить виявляти закономірності БЧР та покращити результати класифікації у різних завданнях.

Для досягнення мети, в роботі поставлено та розв'язано **такі задачі:**

- пошук даних
- підбір архітектури моделі
- тестування розробленої моделі
- розробка інструменту для автоматичного підбору гіперпараметрів моделі.

Об'єкт дослідження: трансформер для класифікації БЧР.

Предмет дослідження: архітектура трансформера для класифікації БЧР.

Наукова новизна роботи: розроблена архітектура моделі трансформера, яка дозволяє виявити різні закономірності в послідовностях.

Практичне значення одержаних результатів. Розроблено інструмент підбору гіперпараметрів моделі, який дозволяє автоматично навчити побудовану модель вирішувати завдання класифікації БЧР для довільних даних.

Апробація. Окремі результати роботи були представлені на наукових конференціях як опубліковані тези:

1. Лісовий В.Г. Архітектура трансформера для класифікації багатовимірних часових рядів // XIV Міжнародна науково-практична конференція молодих учених та студентів «Актуальні задачі сучасних технологій», Тернопіль, 11-12. Грудня 2025 р. с. 289 – 291.

2. Лісовий В.Г., Яцишин В.В., Семенишин Г.М. Алгоритм роботи програмного інструменту автоматичного підбору гіп, ерпараметрів трансформера // Інформаційні моделі, системи та технології: Праці XIII наук.-техн. конф. Тернопіль, 17-18 грудня 2025 р. с. 56 – 57.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Проведення аналізу часових рядів із використанням методів машинного навчання

Часовий ряд є послідовністю числових показників, упорядкованих у часі, які відображають рівень стану і зміни явища, що вивчається [1]. Часові ряди є цінним джерелом інформації для різних галузей людської діяльності. Дослідження часових рядів має важливе значення для розвитку наукових та практичних підходів у різних галузях, таких як промисловість [2], медицина [3], економіка [4], фінанси [5] та інші. Так як часові ряди дозволяють аналізувати і прогнозувати тенденції, виявляти закономірності і розуміти динаміку процесів, що відбуваються.

Аналізування часових рядів прогнозує вилучення важливої зведеної та статистичної інформації з точок даних, заданих у хронологічному порядку [6]. Систематичне і глибоке розуміння часових рядів є невід'ємним виявлення закономірностей, трендів і аномалій. Методи аналізування часових рядів містять широкий спектр технік і методів, розвиток яких дозволяє більш точно моделювати, прогнозувати та інтерпретувати часові ряди, що має важливе значення для прийняття обґрунтованих рішень.

Властиво аналіз часових рядів та їх використання стають все більш значущими в сучасному світі, оскільки ці структури даних набувають все більшого поширення [6]. Тенденції розвитку аналізу часових рядів мають багатогранний характер і відбивають динаміку сучасних технологічних досліджень. В останні роки помітне стрімке зростання зацікавленості до методів власне машинного навчання, насамперед глибокого навчання, що застосовуються в аналізі часових рядів. Так, наприклад, кількість публікацій на тему часових рядів за останні десять років зросла практично вдвічі, як показано на рис. 1.1. Розвиваються інноваційні методи обробки великих даних та виділення прихованих закономірностей у динаміці часу.

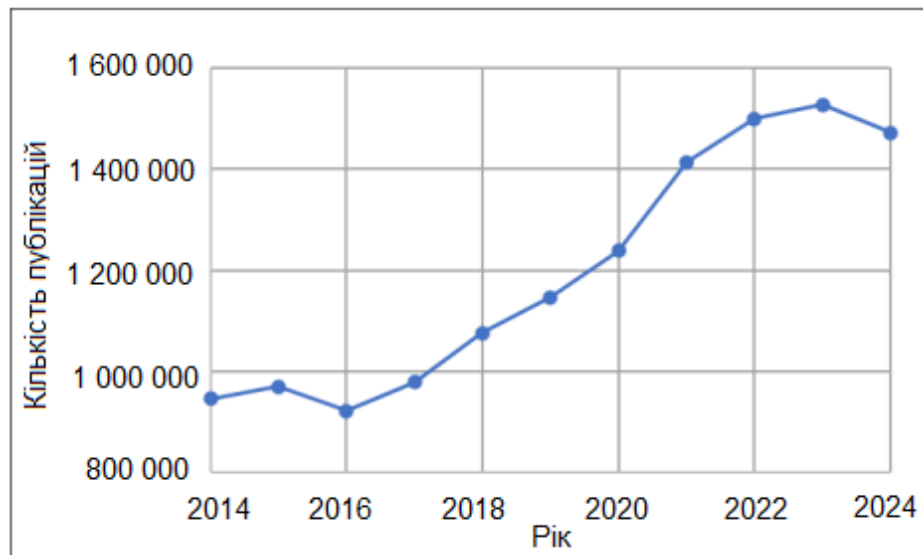


Рисунок 1.1 – Щорічна кількість публікацій з часових рядів

Важлива увага також приділяється технологіям аналізу часових рядів у реальному часі, що має високе значення для прийняття оперативних рішень. У світлі зростаючої складності даних активно досліджуються методи автоматизації передобробки та фільтрації часових рядів. Очікується, що майбутнє аналізу часових рядів буде спрямоване на створення більш ефективних, точних та стійких моделей, здатних адаптуватися до різноманітних сценаріїв використання.

Основні завдання, пов'язані з аналізом часових рядів, включають класифікацію, прогнозування, пошук аномалій і кластеризацію [7].

Класифікація часових рядів полягає у присвоєнні часовому ряду певної категорії чи класу. Необхідно класифікувати часові ряди на основі їхньої поведінки чи характеристик, щоб визначити, чи належить ряд до певного класу.

Пошук аномалій у часових рядах полягає у виявленні незвичайних або аномальних патернів у часових рядах. Аномалії можуть вказувати на непередбачені події чи проблеми у системі.

Кластеризація здатна допомогти детектувати непомітні патерни або групи даних і спростити аналіз часових рядів.

Кожне із завдань аналізу часових рядів має свої специфічні методи та підходи, які вибираються в залежності від характеристик даних та необхідних

результатів. Здебільшого, на вирішення цих завдань застосовуються статистичні методи [8] і методи машинного навчання.

Статистичні методи включають різні підходи, такі як методи згладжування, стаціонарності та оцінки трендів і сезонності. Ці методи засновані на статистичних моделях та теорії, які дозволяють аналізувати та інтерпретувати часові ряди.

Методи машинного навчання, разом із тим, котрі базуються на алгоритмах та моделях, які навчаються на основі історичних даних та можуть автоматично виявляти патерни та залежності у часових рядах.

Застосування методів машинного навчання для аналізу часових рядів досліджувалося як вітчизняними [9], так і зарубіжними вченими [10]. Для вирішення завдань аналізу часових рядів застосовувалися різні методи машинного навчання, зокрема: дерева рішень [11], описові статистики [12], рекурентні [13] та згорткові нейронні мережі [14], а також їх різні комбінації [15]. В даний час найкращих результатів у вирішенні завдань аналізу часових рядів досягають моделі трансформерів [16], які ґрунтуються на механізмі уваги, що враховує контекст послідовності. Ці моделі були спеціально розроблені для аналізу послідовностей, а саме для перекладу тексту з однієї мови на іншу. Трансформери знайшли широке застосування у обробці природної мови (NLP). Ці моделі досягають високих результатів завдяки можливості апроксимувати послідовності [17].

Нині активно розробляються моделі трансформерів на вирішення завдань аналізу часових рядів. Ця тема у науковому співтоваристві є гострою, оскільки вчені сперечаються про ефективність застосування трансформерів до часових рядів. Так, наприклад, у роботах [18, 19, 20] доводиться ефективність застосування трансформерів для прогнозування часових рядів, а в роботі [21] ставиться під сумнів цей доказ шляхом порівняння моделі трансформера з простою повнопов'язаною нейромережею. Результати тестування у роботі вказують на те, що модель трансформера програє в ефективності моделі

нейромережі. У зв'язку з цим наступні дослідження, такі як [22, 23], проводять додаткові тести та порівняння із застосуванням простих нейронних мереж.

Застосування моделей трансформерів до завдань аналізу часових рядів розглянуто у роботі [24], де докладно досліджується історія використання трансформерів у вирішенні даних завдань. Аналізуються найкращі результати, досягнуті у попередніх роботах, із особливостей, що з рішенням завдань аналізу як одновимірних часових рядів, так і БЧР. Ця особливість сильно впливає на архітектуру моделей і стратегію навчання, оскільки багатовимірність дає більше інформації про явище, що вивчається, і накладає обмеження при використанні алгоритмів машинного навчання. Так, наприклад, автори роботи [25] пропонують незвичайний спосіб аналізу БЧР, коли як ембединги приймаються всі точки кожного часового ряду, тоді як у класичних рішеннях використовується по 1 точці з кожного ряду.

Дослідження БЧР привертає дедалі більше уваги в контексті розуміння складних явищ. На відміну від одновимірних рядів, багатовимірні надають більш широку та різнобічну інформацію про динаміку досліджуваного процесу. Це є особливо цінним при вирішенні різних завдань, таких як класифікація, пошук аномалій та кластеризація. Таким чином, аналіз БЧР стає ключовим інструментом для більш глибокого розуміння складних систем та ефективного вирішення різноманітних завдань у галузі науки про дані.

1.2 Застосування моделей трансформерів для класифікації багатовимірних часових рядів

Вирішення завдань класифікації в контексті БЧР наголошує на необхідності урахування взаємозв'язку та варіативності різних вимірів даних. У цьому контексті, для забезпечення якісного навчання та тестування моделей, надзвичайно важливим є вибір відповідного набору даних. Для навчання та тестування моделей БЧР у науковому співтоваристві широко використовується набір даних UEA [26]. UEA є різноманітною колекцією датасетів, що

покривають різноманітні сфери функціонування людини. Кожен датасет включає як навчальну, так і тестувальну частини, надаючи дослідникам надійні дані для навчання та оцінки ефективності моделей БЧР. Крім того, всі часові ряди всередині одного датасета приведені до єдиної розмірності, що забезпечує консистентність та стандартизацію даних. Вибір єдиних даних не тільки забезпечує надійність результатів, а й дозволяє проводити порівняльні аналізи ефективності моделей різних наборів даних. Це важливо для обґрунтованого вибору підходів та моделей, а також для забезпечення можливості повторюваності та зіставлення результатів з аналогічними дослідженнями у науковій спільноті.

Для оцінки ефективності моделей класифікації часових рядів часто застосовується метрика точності (accuracy), що виражає частку правильних відповідей відносно загальної кількості відповідей. Вибір саме цієї метрики обумовлений тим, що всі датасети, що розглядаються, є збалансованими по відношенню до прогнозованих класів. Така властивість забезпечує більш надійну інтерпретацію результатів, оскільки точність враховує як чітко передбачені позитивні, і чітко передбачені негативні випадки. У контексті часових рядів, де важлива як загальна, так і детальна правильність класифікації, використання метрики точності є логічним та інформативним підходом для оцінки якості моделей.

Автори [27] є першими, хто спробував розробити власну архітектуру для класифікації БЧР. У роботі описується архітектура моделі TST, яка складається з послідовності модулів енкодера класичного трансформера. Також у роботі пропонується спочатку передбачити модель прогнозувати пропущені значення у вихідному часовому ряду і після цього навчати модель вирішувати необхідне завдання. Ця стратегія, зазвичай, використовується в NLP [28] для того, щоб модель заздалегідь знала про структуру послідовності та розуміла наявні закономірності. Розроблена архітектура разом із стратегією навчання показали найкращі результати у класифікації БЧР на наборі даних UEA. Для порівняння були протестовані моделі, засновані на класичній описовій статистиці, деревах

рішень, повнозв'язкових згорткових та рекурентних нейромереж. Це була перша модель трансформера, здатна показати більш високі значення метрики якості, ніж усталені алгоритми класифікації БЧР.

У роботі [29] автори спробували розробити власну архітектуру моделі трансформера для класифікації часових рядів – GTN (Gated Transformer Networks). Основна відмінність від TST полягає в тому, що модель складається з 2 великих блоків, один з яких аналізує динаміку даних у часовому ряді, а інший між часовими рядами. Для тестування моделі автори використовують набір даних, представлений у роботі [30], оскільки цей набір даних не піддавався обробці і містить часові ряди з різною розмірністю всередині одного датасета. Отже, цей набір даних є більш наближеним до реальності. Для порівняння розробленої моделі використовують моделі, засновані на повнозв'язкових згорткових та рекурентних нейромережах. Результати показали, що метрика якості у GTN у більшості випадків вища, ніж у аналогів. У роботі не наводиться порівняння з моделлю TST, тому порівняти ці дві архітектури неможливо у огляді літератури.

Грунтуючись на архітектурі TST і ідеї передбачення моделі в роботі [31] пропонується нова архітектура - TARNet (Task-Aware Reconstruction). Ідея даної архітектури ґрунтується на тому, що модель у процесі навчання автоматично вчиться визначати «важливість» підпоследовностей за допомогою карток уваги і в той же час прогнозувати лише «важливі» пропущені значення. Ця архітектура і стратегія навчання під час тестування показали найкращі результати. Порівняння проводилося з алгоритмами класичної описової статистики, дерев рішень, повнозв'язкових згорткових та рекурентних нейромереж, а також з TST. Основним набором даних для навчання та тестування був UEA, але також були використані інші. Тестування показало, що точність моделі TARNet вища у більшості випадків, ніж у аналогів при класифікації БЧР.

Найактуальнішою роботою на момент написання роботи є [32]. Автори розробили свою архітектуру моделі – Dyformer (dynamic transformer). Ключова ідея моделі полягає в автоматичному доборі структури трансформера для

вирішення завдань. Ця модель складається з множини модулів, у кожному з яких автоматично регулюється додавання розрахункового блоку. Також автори роботи є першими, хто спробував провести автоматичну декомпозицію часових рядів для завдання класифікації. Тестування проводилося на наборі даних UEA шляхом порівняння Dyformer з іншими моделями, що включають алгоритми класичної описової статистики, повнозв'язкових згорткових і рекурентних нейромереж, а також моделі трансформерів TST, TARNet. Результати показали, що метрика якості на тестовій вибірці для класифікації багатовимірних рядів у Dyformer вище, ніж у аналогів. Тому на даний момент Dyformer вважається SOTA -моделлю.

Застосування моделей трансформерів для класифікації часових рядів є активним напрямом досліджень, що знаходить підтвердження у низці сучасних робіт [33, 34]. Зокрема, у роботі [33] здійснюється успішне застосування основного алгоритму моделей трансформерів – механізму уваги – для ефективної обробки звукових даних. Інший приклад такого застосування представлений у роботі [34], де застосовується модель класичного трансформера для класифікації супутникових даних.

1.3 Стратегії токенизації

Під токенизацією розуміється процес поділу вихідної послідовності на базові одиниці, які потім аналізуються щодо взаємозв'язків [35]. Після токенизації виходять токени, які можуть бути словами, підсловами, точками та інші смисловими одиницями. Цей процес необхідний для представлення послідовності як абстрактних смислових одиниць для подальшого аналізу алгоритмами.

При обробці БЧР використовується одна із трьох основних стратегій токенизації.

1.3.1 TPT

Стратегія TPT ролягає у тому, що як токен приймаються крапки в БЧР у кожен час, як показано на рис. 1.2. Цей підхід дозволяє шукати взаємозв'язки між станами системи протягом часу і застосовується в багатьох розв'язках задач аналізу БЧР [18, 19].

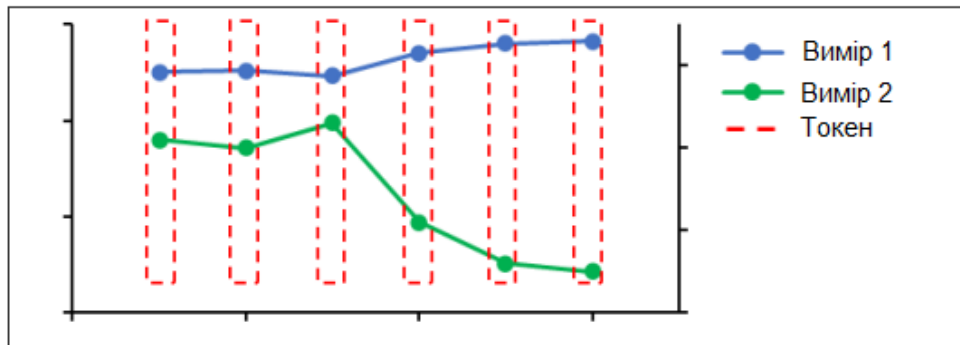


Рисунок 1.2 – TPT токенизація

1.3.2 TVT

Ідея стратегії TVT у тому, що як токен беруться крапки в БЧР у кожному вимірі, як показано на рис. 1.3. Цей підхід дозволяє шукати взаємозв'язки між станами системи між вимірами та використовується в роботі [25].

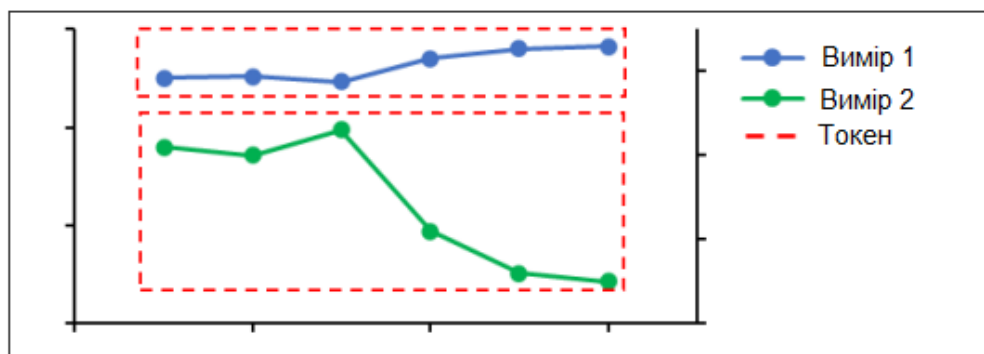


Рисунок 1.3 – TVT токенизація

1.3.3 Patching

Стратегія Patching полягає в тому, що дані поділяються на

підпоследовності однакової розмірності (патчі) і за токен приймаються точки в цих підпоследовностях, як показано на рис. 1.4. Цей підхід дозволяє шукати взаємозв'язки між станами системи в локальних областях протягом часу вздовж кожного виміру та використовується в роботі [22].

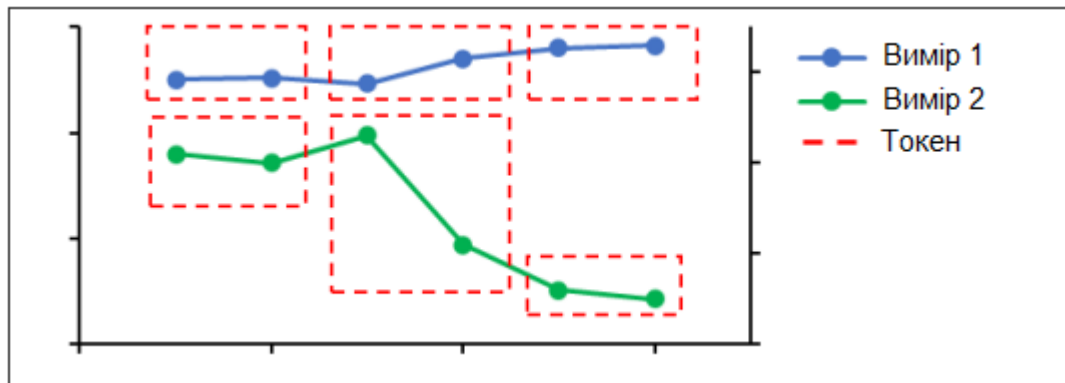


Рисунок 1.4 – Patching токенизація

1.4 Недоліки застосування існуючих моделей трансформерів для класифікації багатовимірних часових рядів

Для аналізу послідовностей моделі трансформерів використовують механізм уваги [16]. Цей алгоритм дозволяє шукати взаємозв'язки між довільними послідовностями. Цей механізм дозволяє моделі звертати увагу на різні частини вхідних послідовностей залежно від їхньої значущості для конкретного завдання аналізу або обробки даних. Шляхом навчання ваги модель може визначити, які елементи послідовності слід враховувати при виконанні конкретної задачі, що робить її більш ефективною в аналізі та використанні інформації.

Нині жоден із підходів до вирішення завдання класифікації БЧР не враховує локальні взаємозв'язки між вимірами протягом часу. Використовуючи існуючі стратегії токенизації, розроблені моделі не враховують локальні зміни вимірів один щодо одного, як показано на рис. 1.5.

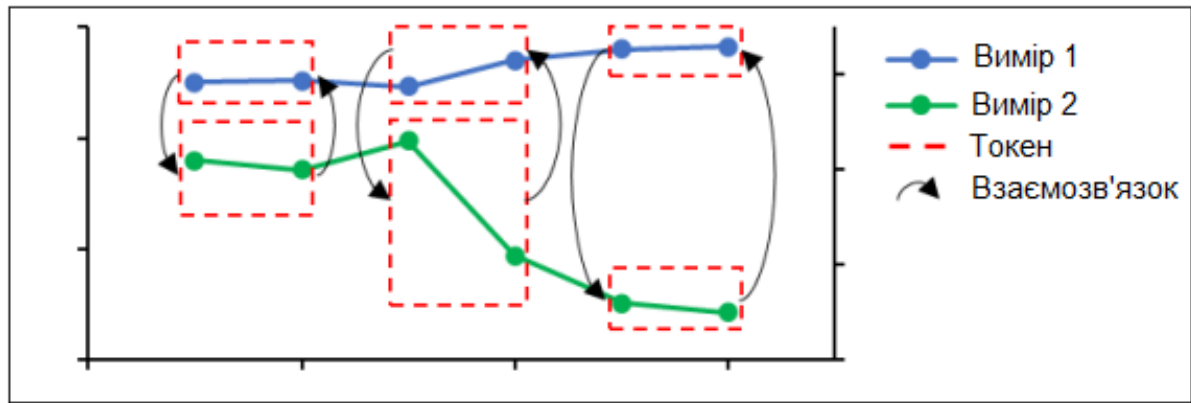


Рисунок 1.5 – Взаємозв'язок між локальними областями серед вимірів

Відсутність пошуку взаємозв'язків між локальними областями серед вимірів при розв'язанні задачі класифікації БЧР веде до втрати інформації про залежність між локальними змінами, які можуть характеризувати той чи інший клас.

Існуючі моделі трансформерів прагнуть покращити результати шляхом різних стратегій, таких як ускладнення самих моделей чи алгоритмів передобробки даних. Наприклад, роботи [27, 31, 32] використовують навчання шляхом вирішення задачі маскування, що дозволяє моделям отримати більш загальні уявлення про дані. В інших дослідженнях, наприклад, у роботі [32] застосовується декомпозиція часових рядів з метою покращення результатів класифікації. Однак, такі підходи суттєво ускладнюють донавчання та застосування моделей трансформерів до БЧР. Це зумовлено потребою налаштування значного числа параметрів чи виконання додаткових кроків передобробки даних, що потребує додаткових обчислювальних ресурсів та часу.

У зв'язку з цим потрібна розробка моделі класифікації БЧР, яка здатна враховувати локальні взаємозв'язки між різними вимірами та легковажну архітектуру. Це дасть змогу покращити ефективність класифікації даних, мінімізуючи складність моделі та знижуючи обчислювальні витрати, пов'язані з її навчанням та застосуванням.

1.5 Висновки до першого розділу

У цьому розділі був проведений ґрунтовний аналітичний огляд літературних джерел, які стосуються особливостей використання моделей трансформерів для аналізу часових рядів.

Наведено специфіку проведення аналізу часових рядів при застосуванні методів машинного навчання. Розглянуто використання моделей трансформерів для класифікації БЧР. Описані три основні моделі токенізації (TPT, TVT, Patching). Наведено недоліки наявних моделей трансформерів, котрі застосовуються для класифікації БЧР.

2 ТЕОРЕТИЧНА ЧАСТИНА

2.1 Архітектура розроблюваної моделі

Для вирішення задачі класифікації БЧР розроблено архітектуру моделі трансформера, що дозволяє враховувати як взаємозв'язки між локальними областями серед вимірів, так і взаємозв'язки між підпоследовностями вздовж кожного виміру. Як основа взята архітектура двобаштової моделі [36], яка може обробляти різну природу даних. Розроблена архітектура показано на рис. 2.1.

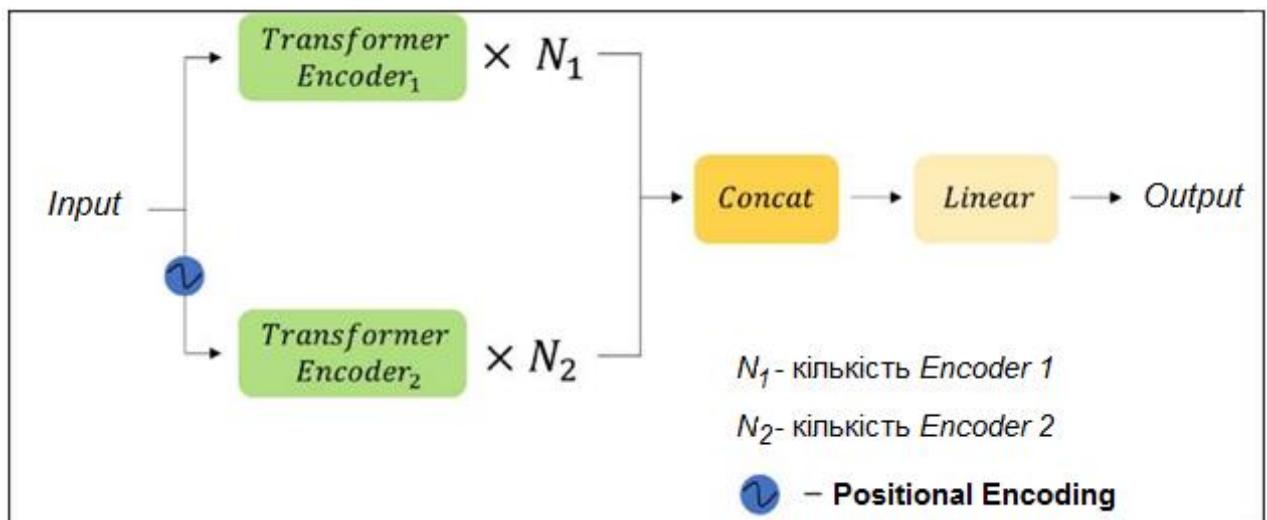


Рисунок 2.1 – Архітектура моделі, що розробляється

Вхідні дані надходять у два різних кодувальники, званих Encoder 1 і Encoder 2, додаючи у другому випадку позиційний вектор. Далі дані проходять через N шарів кодувальників, що об'єднуються, проходять через лінійний шар і формується цільовий результат [37].

Як вхідні дані використовуються БЧР, розділені на підпоследовності однакової розмірності (патчі), як показано на рис. 2.2. При проходженні даних через розроблену модель формується набір ймовірностей, що дозволяє визначити клас найбільшої ймовірності.

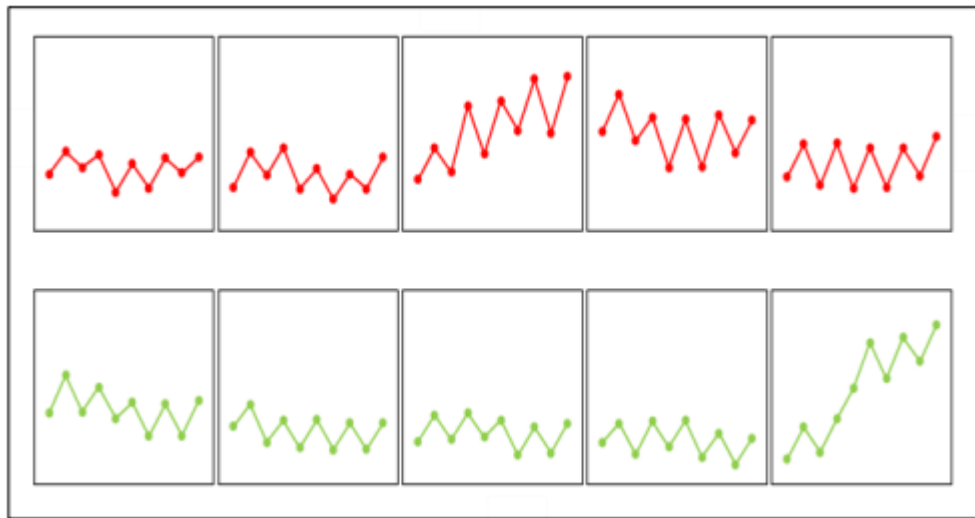


Рисунок 2.2 – Вихідні дані

Як кодувальники використовуються оригінальні шари трансформер моделі, представленої в роботі [16]. Дані шари перетворюють вхідну послідовність у внутрішнє представлення, що містить інформацію про семантику та структуру вихідних даних. Він складається з декількох блоків, кожен з яких містить безліч шарів самоуваги (self-attention) і шарів прямого (feed-forward). Перші шари допомагають моделі фокусуватися на різних частинах вхідних даних, а другі шари забезпечують нелінійні перетворення, покращуючи здатність моделі до вивчення складних залежностей даних. Кількість кожного кодувальника може змінюватись і є гіперпараметром моделі. Завданням Encoder 1 є пошук закономірностей між патчами серед вимірів. А завданням Encoder 2 є пошук закономірностей між патчами вздовж кожного часового ряду. Пошук взаємозв'язків кодувальниками показано на рис. 2.3.

Основна ідея даної архітектури полягає у пошуку закономірностей між вимірами, що дозволяє досягти Encoder 1. Оскільки кодувальники не враховують упорядкованість послідовностей, прийнято додавати позиційний вектор до даних. Але оскільки серед вимірів немає порядку, цей вектор додається тільки до даних, які проходять через Encoder 2.

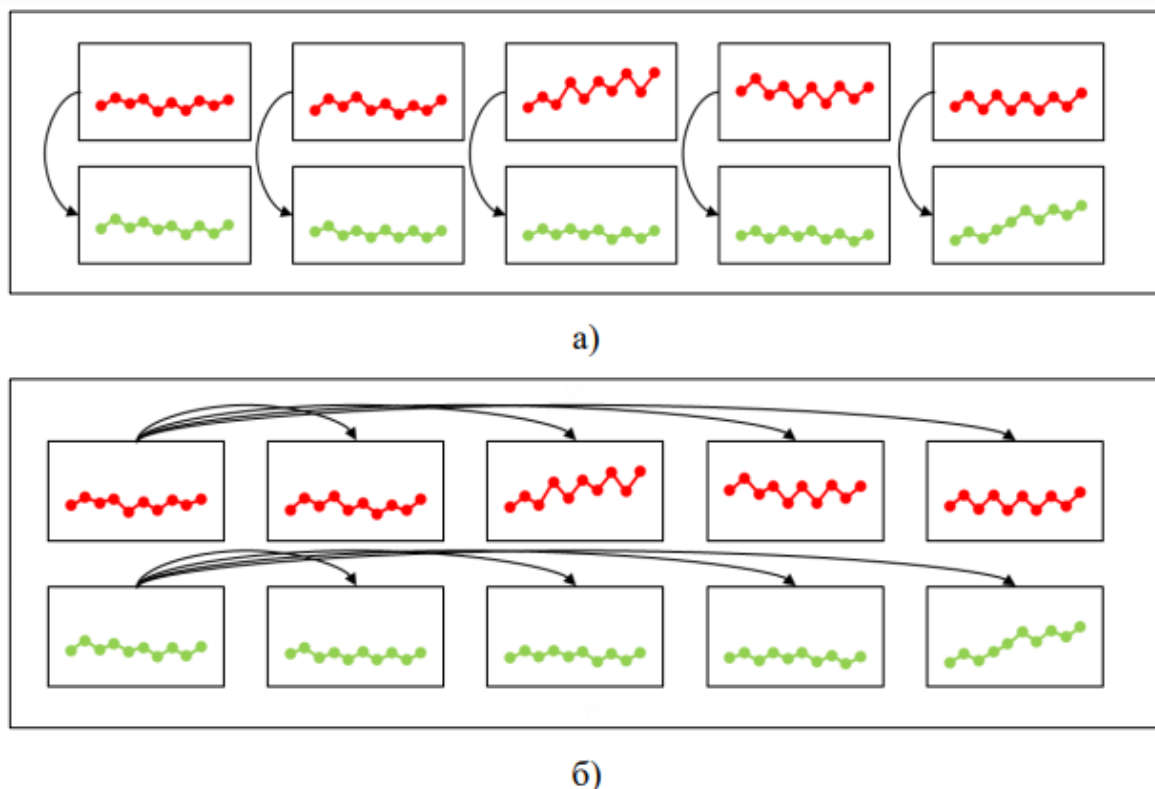


Рисунок 2.3 – Пошук закономірностей: а) Encoder 1; б) Encoder2

Як стратегія позиціонування використовується оригінальний метод з роботи [16], який полягає у використанні тригонометричних функцій для отримання позиційного вектора. Позиційне кодування дозволяє враховувати абсолютні та відносні позиції патчів вздовж кожного виміру.

2.2 Опис даних для оцінки розробленої моделі

Для оцінки розробленої архітектури завдання класифікації БЧР було встановлено критерії відбору даних. Вибрані дані повинні представляти різні сфери та мати різноманітну природу. Крім того, вони повинні мати різні розміри вхідних і вихідних даних, а також бути актуальними, зведеними до однієї довжини і попередньо очищеними від шумів і аномалій.

Під ці критерії підійшов набір даних UEA [26]. Цей набір був зібраний у 2018 році університетом східної Англії та містить 30 датасетів із різних сфер діяльності людини, включаючи медицину, спорт, астрономію та інші. Короткий звіт по всіх датасетах представлений у таблиці 2.1.

Таблиця 2.1 - Зведена таблиця за всіма датасетами

Датасет	Розмір тренувальної вибірки	Розмір тестової вибірки	Кількість вимірів	Довжина прикладу	Кількість класів
ArticulatoryWord Recognition	275	300	9	144	25
AtrialFibrillation	15	15	2	640	3
BasicMotions	40	40	6	100	4
Character Trajectories	1422	1436	3	182	20
Cricket	108	72	6	1197	12
DuckDuckGeese	50	50	1345	270	5
EigenWorms	128	131	6	17984	5
Epilepsy	137	138	3	206	4
Ethanol Concentration	261	263	3	1751	4
ERing	30	270	4	65	6
FaceDetection	5890	3524	144	62	2
FingerMovements	316	100	28	50	2
HandMovement Direction	160	74	10	400	4
Handwriting	150	850	3	152	26
Heartbeat	204	205	61	405	2
InsectWingbeat	30000	20000	200	30	10
JapaneseVowels	270	370	12	29	9
Libras	180	180	2	45	15
LSST	2459	2466	6	36	14
MotorImagery	278	100	64	3000	2
NATOPS	180	180	24	51	6
PenDigits	7494	3498	2	8	10
PEMS-SF	267	173	963	144	7
Phoneme	3315	3353	11	217	39
RacketSports	151	152	6	30	4
SelfRegulation SCP1	268	293	6	896	2
SelfRegulation SCP2	200	180	7	1152	2
SpokenArabic Digits	6599	2199	13	93	10
StandWalkJump	12	15	4	2500	3
UwaveGesture Library	120	320	3	315	8

З метою глибшого висвітлення результатів тестування розробленої архітектури було проведено ретельний аналіз кожного з датасетів.

ArticulatoryWordRecognition. Датасет з визначення слів, що вимовляються, за допомогою електромагнітного артикулографа. Давачі приладу кріпляться на лоб, губи та щелепу людини. Далі вони записують зміни у часі за трьома осями з частотою дискретизації 200 Гц. Приклад датасету показано на рис. 2.4.

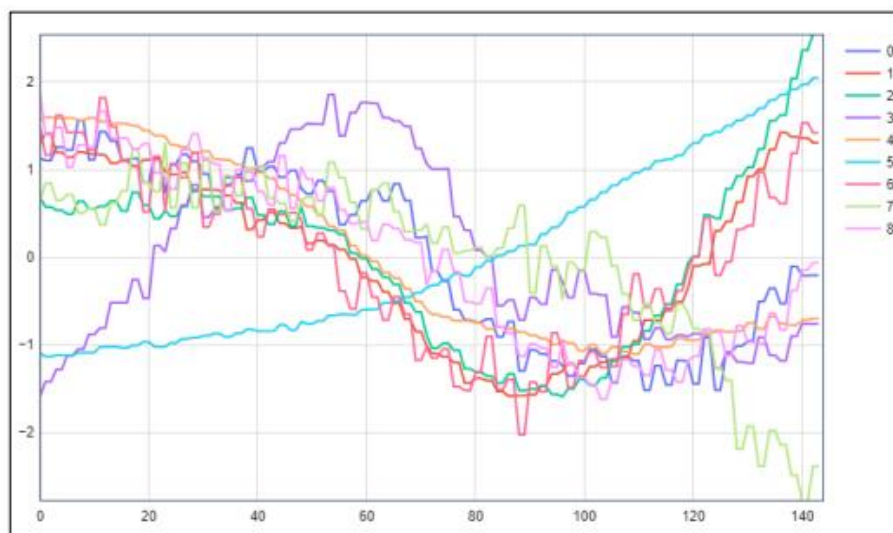


Рисунок 2.4 – Приклад датасету ArticulatoryWordRecognition

AtrialFibrillation. Датасет щодо визначення типів артеріальних фібриляцій передсердь, записаних за допомогою електрокардіографії з різними каналами. Для визначення фібриляції передсердь запис ведеться упродовж 5 секунд. Приклад датасету показано на рис. 2.5.

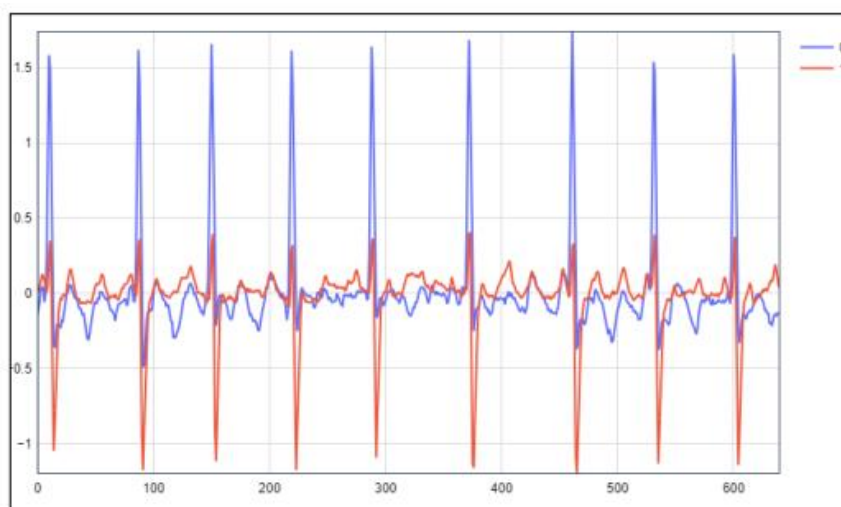


Рисунок 2.5 – Приклад датасету AtrialFibrillation

BasicMotions. Датасет щодо визначення рухів людини з використанням даних з акселерометра та гіроскопа. Давачі були прикріплені до тіла людини та записували зміни у часі за трьома осями. Приклад датасету показано на рис 2.6.

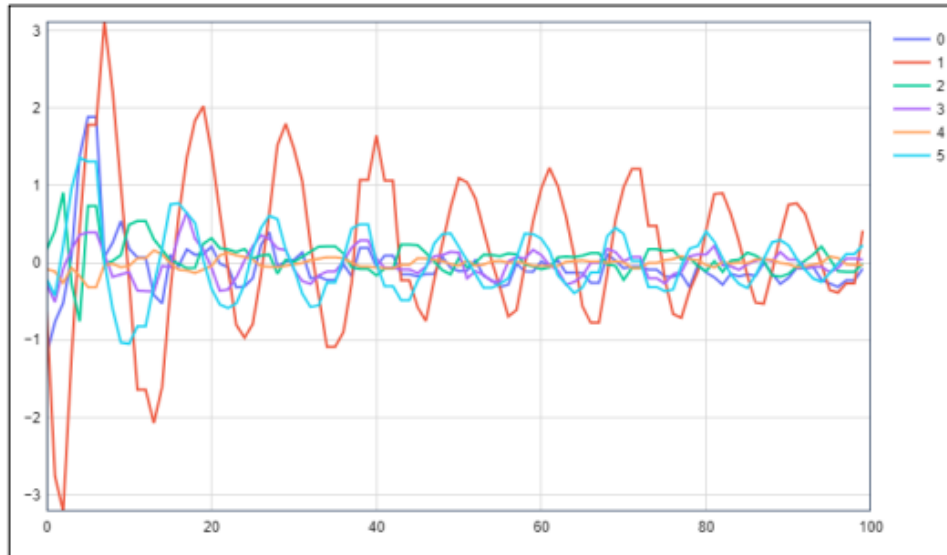


Рисунок 2.6 – Приклад датасету BasicMotions

CharacterTrajectories. Датасет з визначення рукописних цифр з їхньої траєкторії написання. Траєкторія записувалася спеціальним планшетом, у якому зберігалися координати по двох осях і ступінь натискання на ручку людиною. Частота дискретизації 200 Гц. Приклад датасету – рис. 2.7.

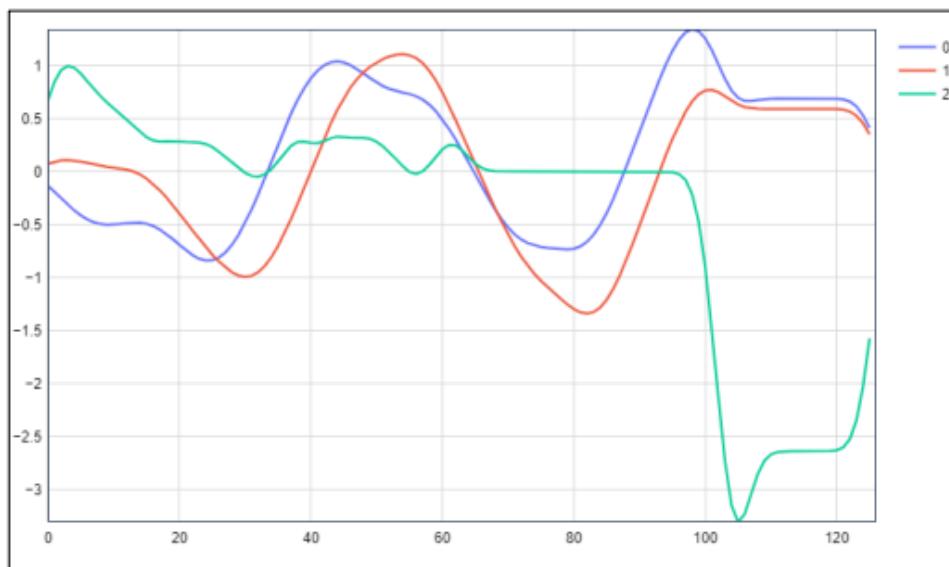


Рисунок 2.7 – Приклад датасету CharacterTrajectories

Cricket. Датасет щодо визначення жестів судді у грі крикет. Дані записувалися за допомогою акселерометрів на руках судді із частотою дискретизації 184 Гц. Акселерометри записували зміни у часі по трьох осях. Приклад датасету показано на рис. 2.8.

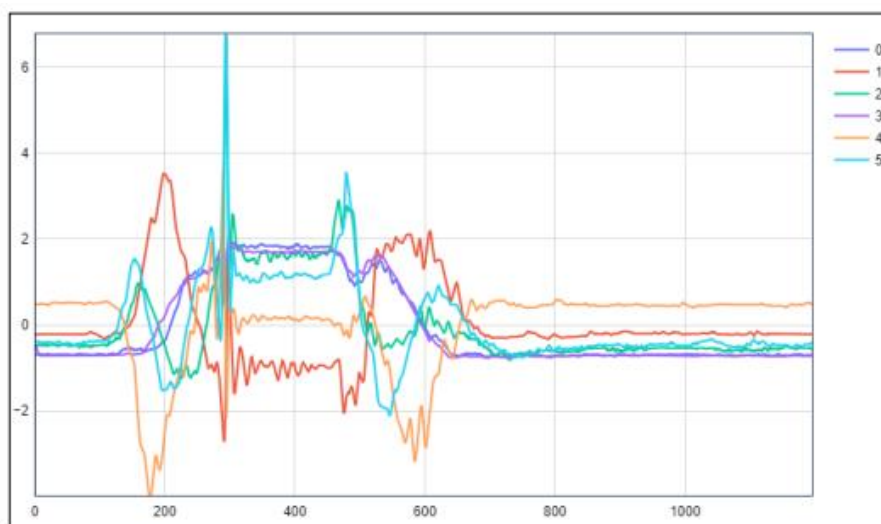


Рисунок 2.8 – Приклад датасету Cricket

DuckDuckGeese. Датасет з визначення породи гусей за звуками, що видаються. Кожен приклад має одну із двох категорій якості. Запис звуку було оброблено та збережено у вигляді спектрограми з частотою 44100 Гц. Приклад датасету показано на рис. 2.9.

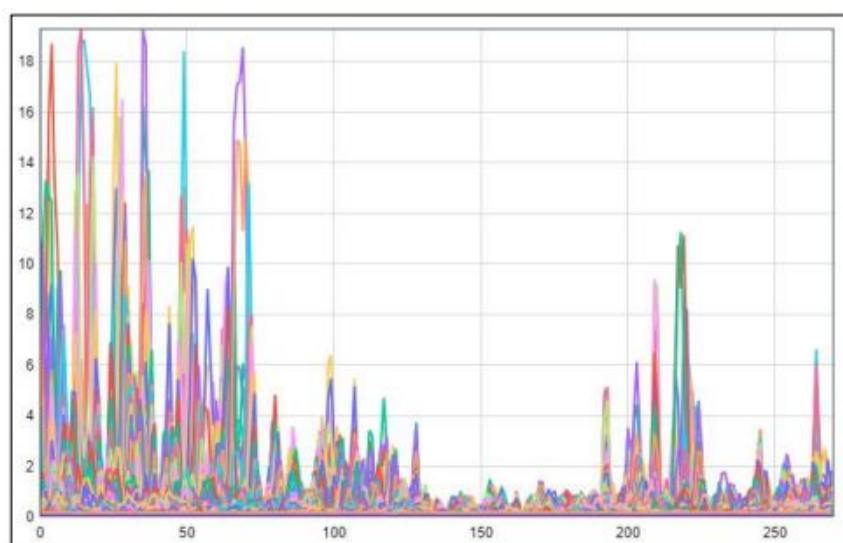


Рисунок 2.9 – Приклад датасету DuckDuckGeese

EigenWorms. Датасет за визначенням сімейства черв'яків за їх рухом. Рух записується людиною як амплітуда відхилення від прямої у трьох вимірах. Класифікація проводиться з урахуванням унікальності руху черв'яків. Приклад датасету показано на рис. 2.10.

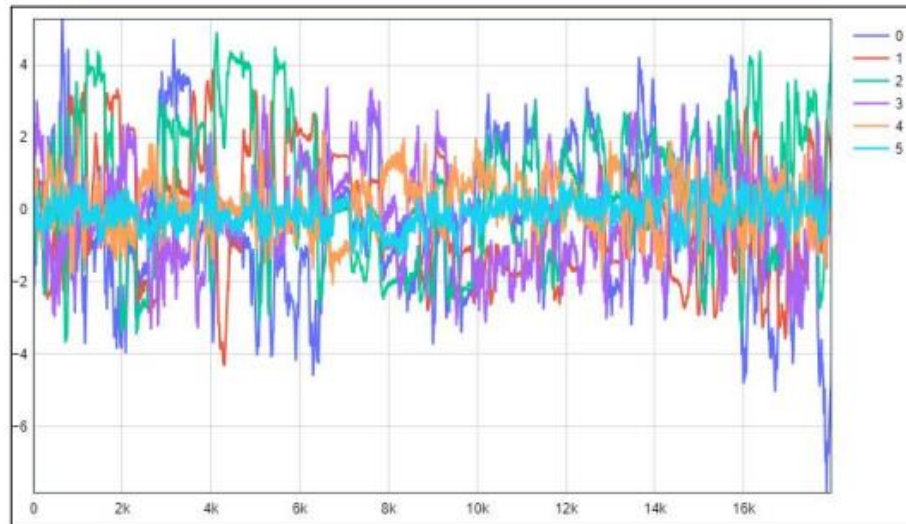


Рисунок 2.10 – Приклад датасету EigenWorms

Epilepsy. Датасет щодо визначення стану людини. Стан визначається за допомогою акселерометра на домінуючому зап'ясті людини імітує різні стани. Акселерометр записував зміни у часі за трьома осями з частотою дискретизації 16 Гц. Приклад датасету показано на рис. 2.11.

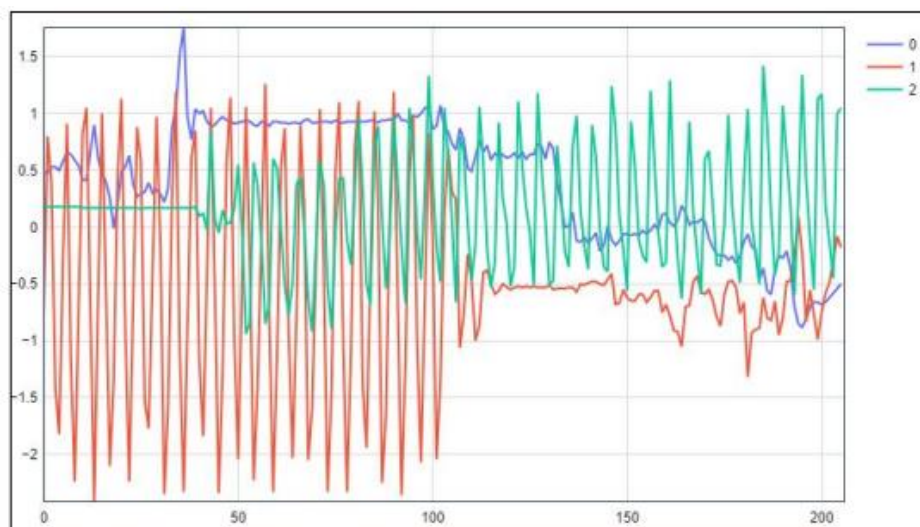


Рисунок 2.11 – Приклад датасету Epilepsy

Ering. Датасет визначення жестів рук за допомогою спеціального приладу у вигляді кільця, що використовує електричне поле. Кожен приклад є вимірюванням за допомогою електрода, який варіюється в залежності від відстані до руки. Приклад датасету показано на рис. 2.12.

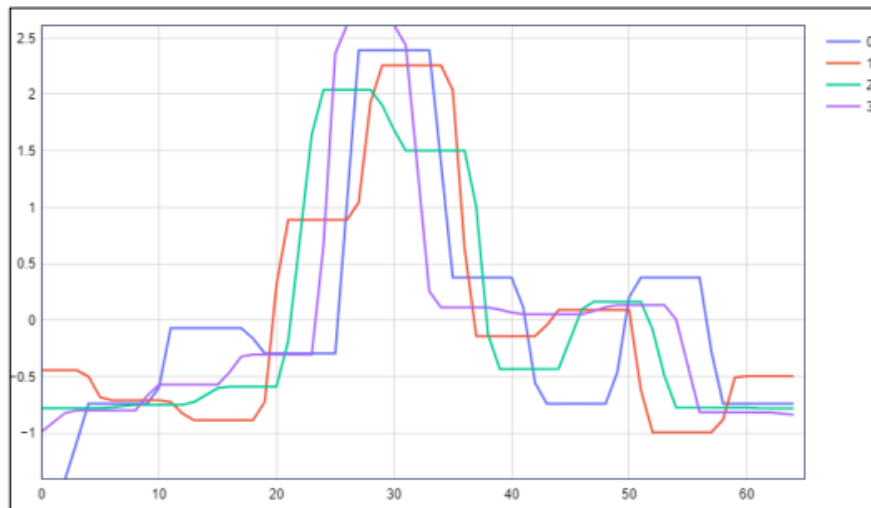


Рисунок 2.12 – Приклад датасету Ering

EthanolConcentration. Датасет визначення концентрації етанолу в розчинах. Кожен показ полягає в тому, що ємність з розчином поміщають між джерелом світла та спектроскопом і далі зберігають спектри різних довжин хвиль. Приклад датасету показано на рис. 2.13.

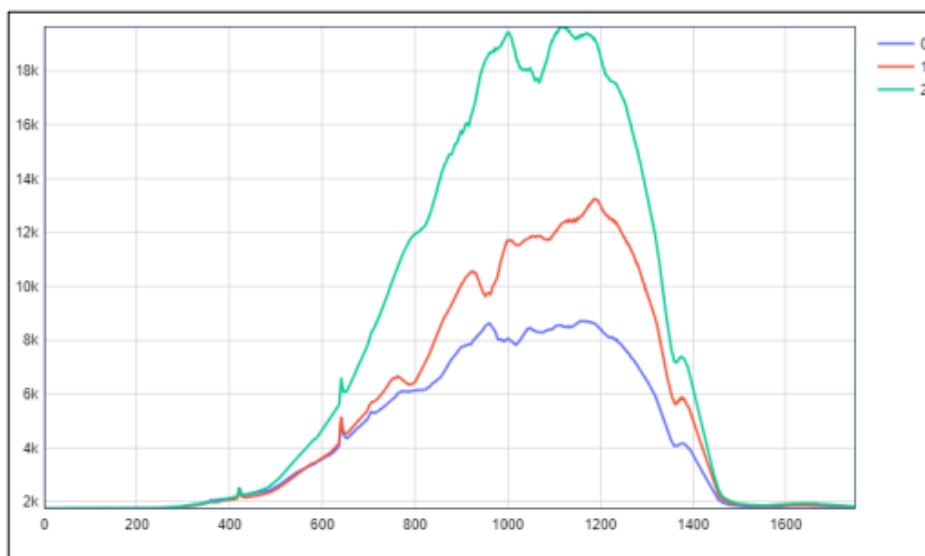


Рисунок 2.13 – Приклад датасету EthanolConcentration

FaceDetection. Датасет визначення знайомих людині осіб за допомогою магнітоенцефалографії. Даний метод показує магнітні поля головного мозку та зберігає їх як БЧР. Приклад датасету показано на рис 2.14.

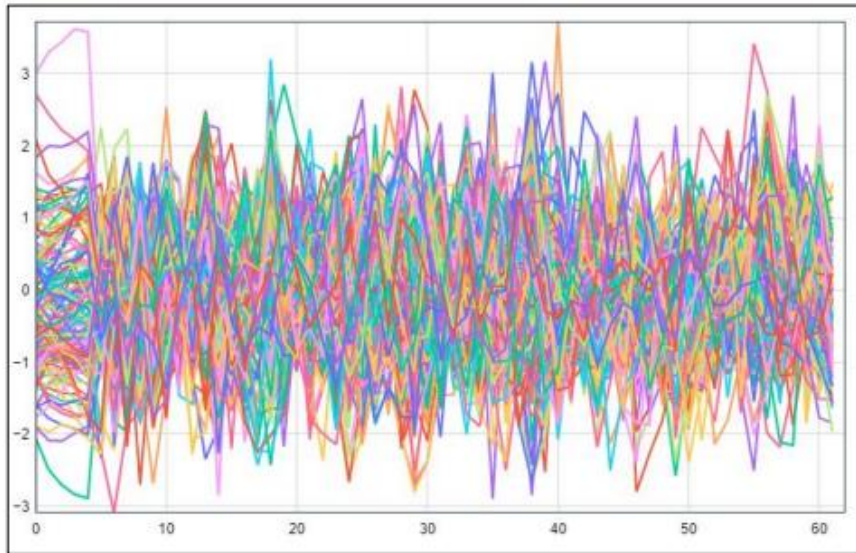


Рисунок 2.14 – Приклад датасету FaceDetection

FingerMovements. Датасет щодо визначення рухів пальців рук за електроенцефалограмою. Даний метод записує електричні поля головного мозку та дозволяє визначити палець якої руки був використаний людиною. Приклад датасету показано на рис 2.15.

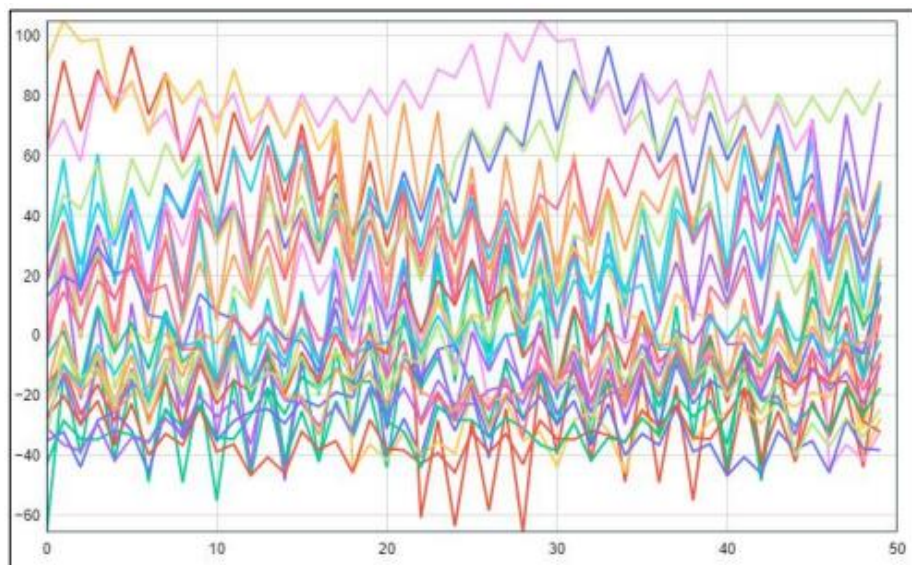


Рисунок 2.15 – Приклад датасету FingerMovements

HandMovementDirection. Датасет щодо визначення напрямку руху руки по магнітоенцефалографії. Давачі зберігають зміни магнітних полів головного мозку та зберігають їх як БЧР. Приклад датасету показано на малюнку 2.16.

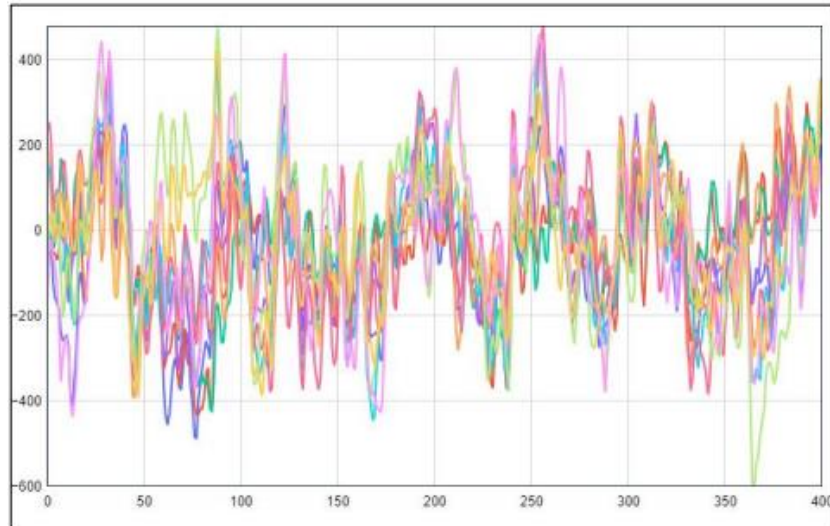


Рисунок 2.16 – Приклад датасету HandMovementDirection

HandWriting. Датасет для визначення написаної людиною літери за допомогою акселерометрів, прикріплених до руки. Акселерометри записують зміни координат по трьох осях у часі. Приклад датасету показано на рис. 2.17.

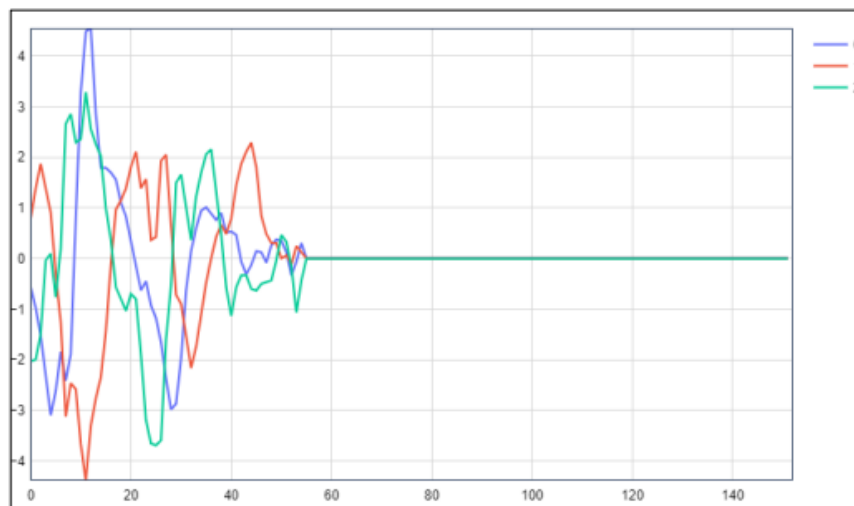


Рисунок 2.17 – Приклад датасету HandMovementDirection

HeartBeat. Датасет визначення хворої людини за її серцебиттям. Ритм серця записувався за допомогою мікрофонів на різних ділянках тіла людини та

результат зберігався у вигляді спектрограми. Приклад датасету показано на рис. 2.18.

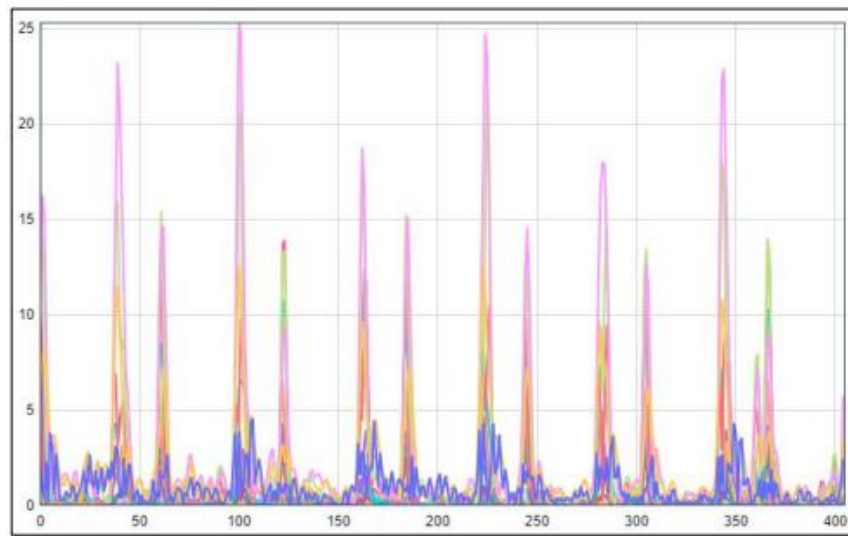


Рисунок 2.18 – Приклад датасету HeartBeat

JapaneseVowels. Датасет з ідентифікації людини з вимови голосних букв. Дані були записані за допомогою мікрофона під час вимови людиною японських букв «а» та «е». Приклад датасету наведено на рис. 2.19.

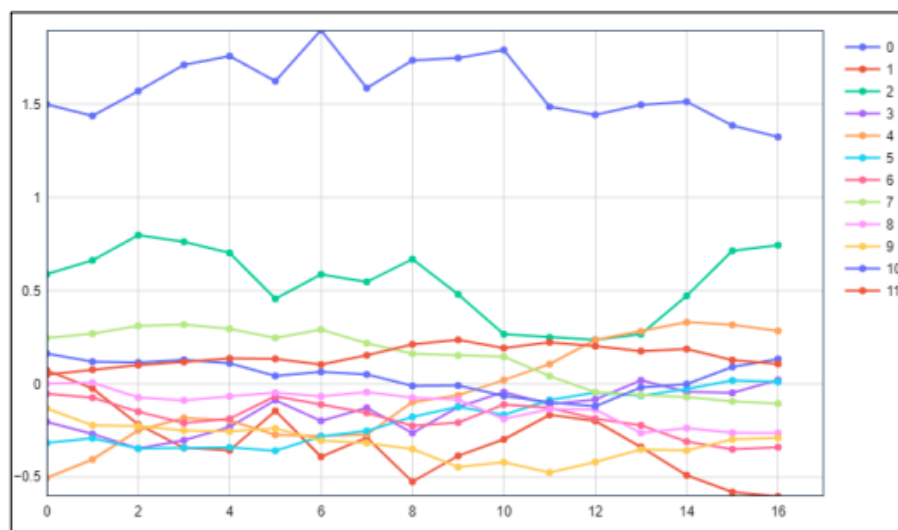


Рисунок 2.19 – Приклад датасету JapaneseVowels

Libras. Датасет визначення бразильської мови жестів. Дані були отримані з відеозаписів, де були показані літери. Відеозаписи були оброблені та записані

у вигляді зміни положення рук у просторі. Приклад датасету показано на рис. 2.20.

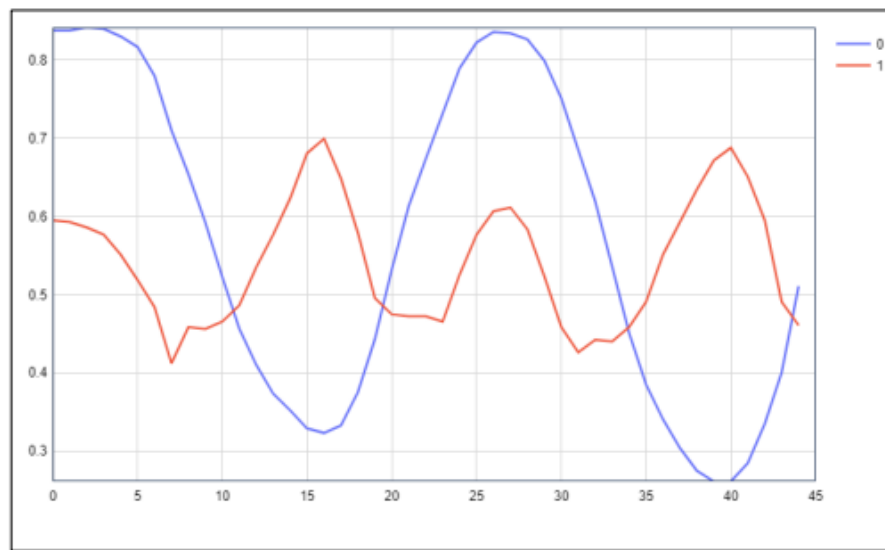


Рисунок 2.20 – Приклад датасету Libras

LSST. Датасет визначення астрономічних об'єктів за допомогою даних, отриманих з телескопа LSST. Дані характеризують зміни світла від об'єктів у різних спектрах, включаючи ультрафіолетове, оптичне та інфрачервоне випромінювання. Приклад датасету представлено на рис 2.21.

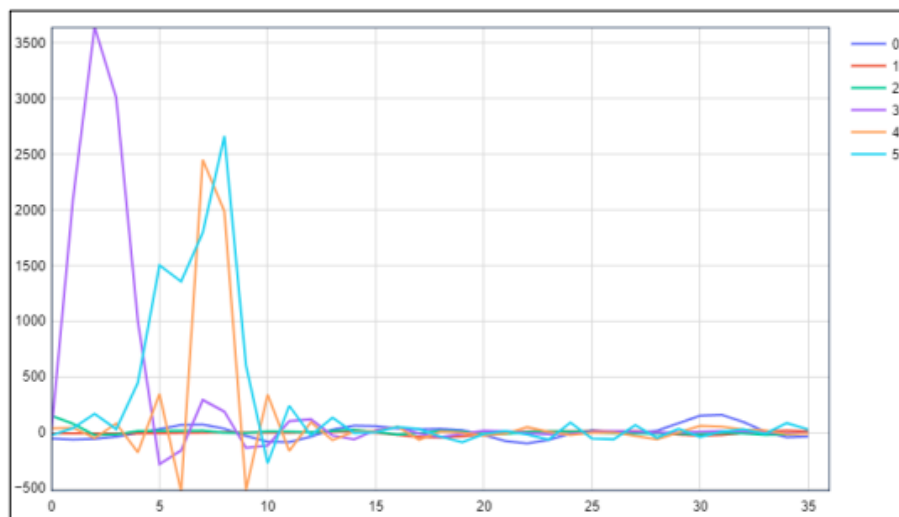


Рисунок 2.21 – Приклад датасету LSST

MotorImagery. Датасет щодо визначення руху пальця або мови з електрокортикографії. Дані характеризують електричну активність мозку під час руху різними ділянками тіла людини. Приклад датасета подано на рис. 2.22.

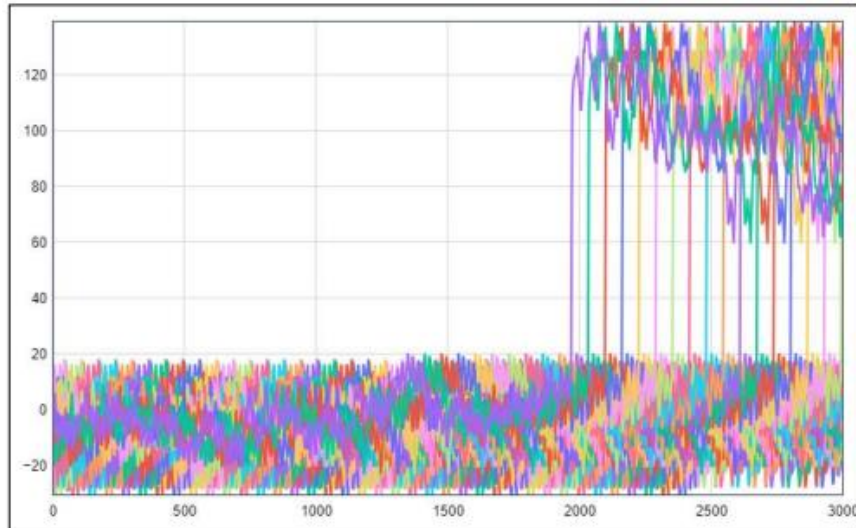


Рисунок 2.22 – Приклад датасету MotorImagery

NATOPS. Датасет для визначення жестів, які показує людина пілотам літаків на злітній смузі. Дані були отримані з плечей, ліктів, зап'ясть і пальців випробуваного у трьох координатах. Приклад датасету показано на рис. 2.23.

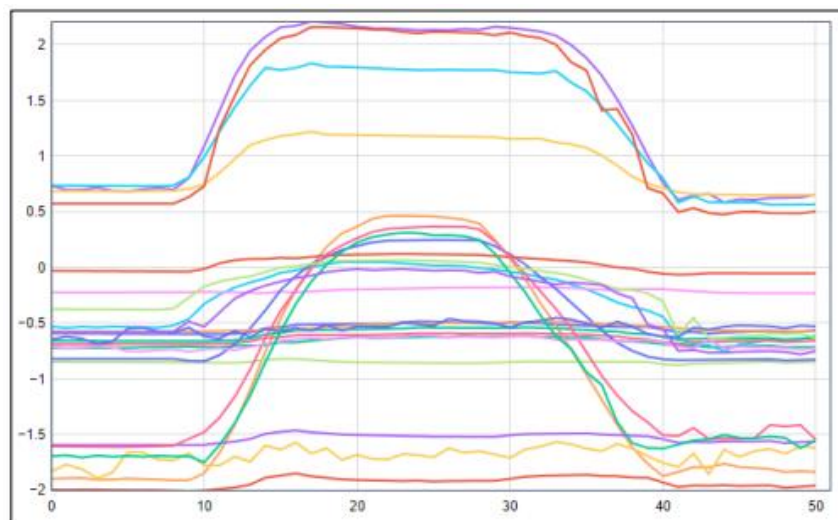


Рисунок 2.23 – Приклад датасету NATOPS

PEMS-SF. Датасет по визначенню днів тижня транспортним потоком у різних ділянках Сан-Франциско. Дані записувалися на спеціальних станціях із певним проміжком часу. Приклад датасету представлено на рис. 2.24.

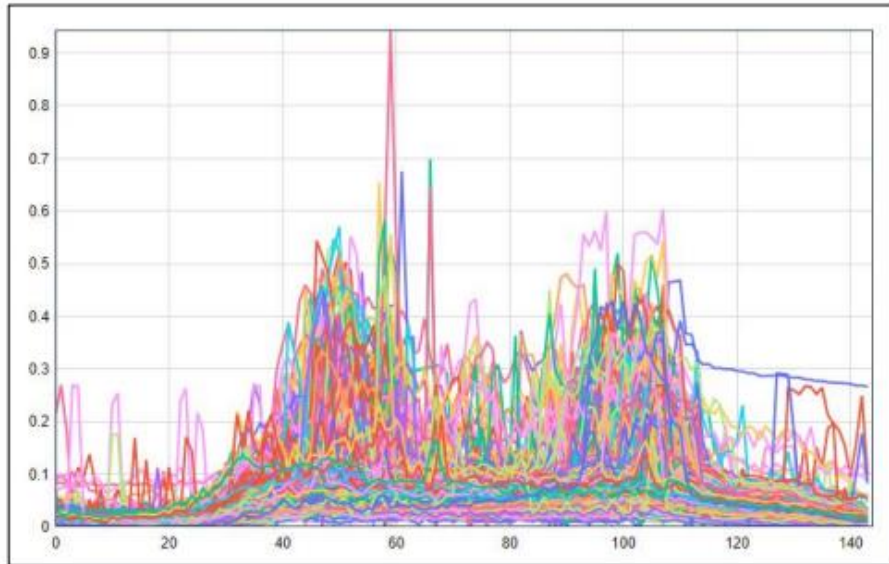


Рисунок 2.24 – Приклад датасету PEMS-SF

PenDigits. Датасет з визначення рукописних цифр за послідовністю двомірних координат ручки, котра використовується. Дані були збережені у високій роздільній здатності, а далі перетворені на нижчу і відсортовані в часі. Приклад датасета подано на рис. 2.25.

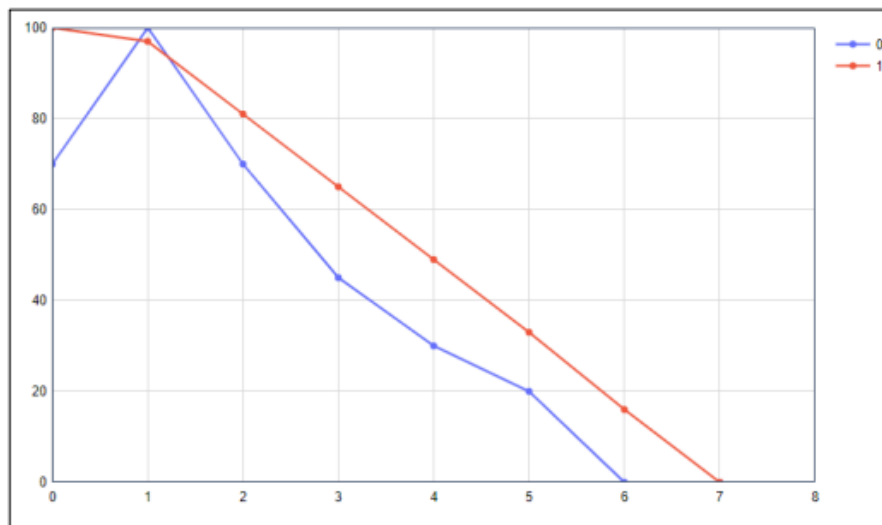


Рисунок 2.25 – Приклад датасету PenDigits

Phoneme. Датасет для визначення фонем, які вимовляються чоловіками та жінками, англійською мовою. Дані були зібрані з аудіодоріжок, попередньо оброблені та збережені у вигляді спектрограм. Приклад датасету показано на рис 2.26.

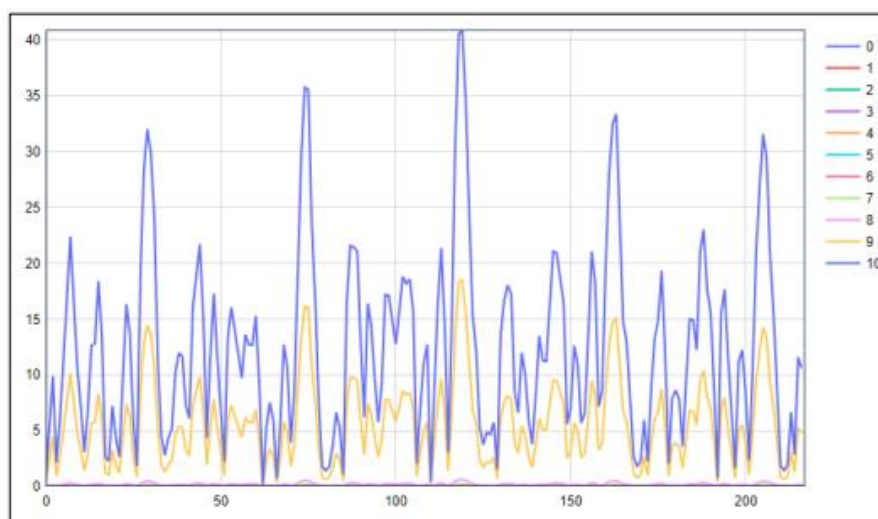


Рисунок 2.26 – Приклад датасету Phoneme

RacketSports. Датасет з визначення виду спорту та типу удару. Випробувані гри у бадмінтон та сквош. Дані з акселерометрів та гіроскопів записувалися з годинника на руках гравців. Збережені результати є змінами параметрів у трьох осях. Приклад датасету показано на рис. 2.27.

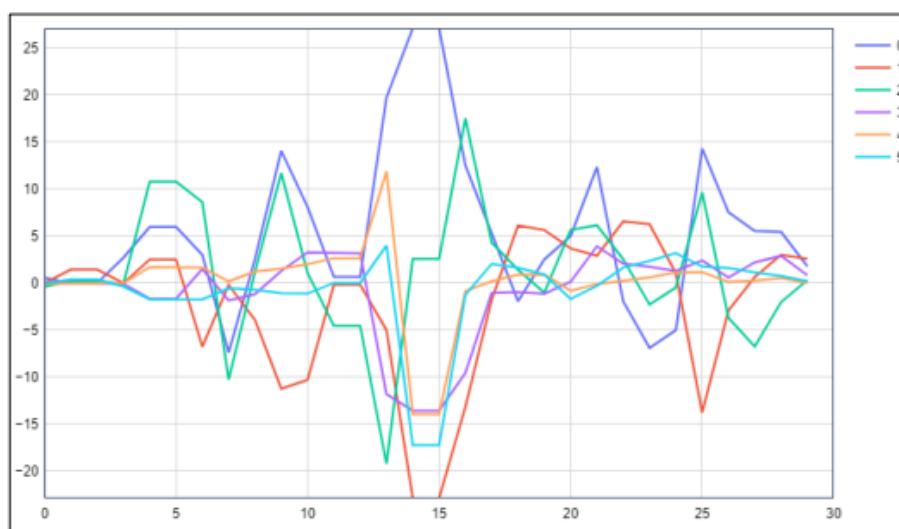


Рисунок 2.27 – Приклад датасету RacketSports

SelfRegulationSCP1, SelfRegulationSCP2. Датасети щодо визначення руху курсору з мозкової активності людини, записаної за допомогою електроенцефалограми. Для визначення активності використовувався параметр коркових потенціалів. Приклад датасетів показано на рис. 2.28.

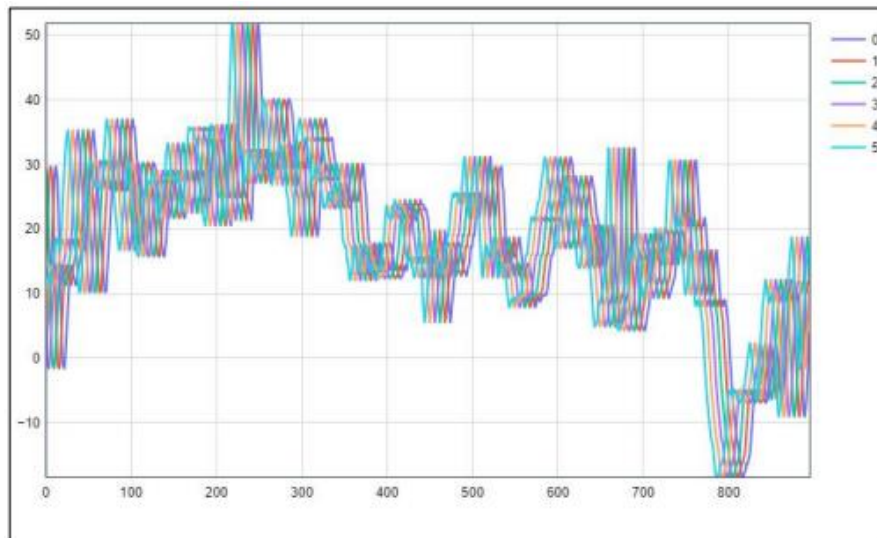


Рисунок 2.28 – Приклад датасету SelfRegulationSC

SpokenArabicDigits. Датасет за визначенням арабських літер за вимовою. Дані є частотними кепстральними коефіцієнтами (MPCC), які характеризують короткочасний спектр потужності звуку. Приклад датасету показано на рис. 2.29.

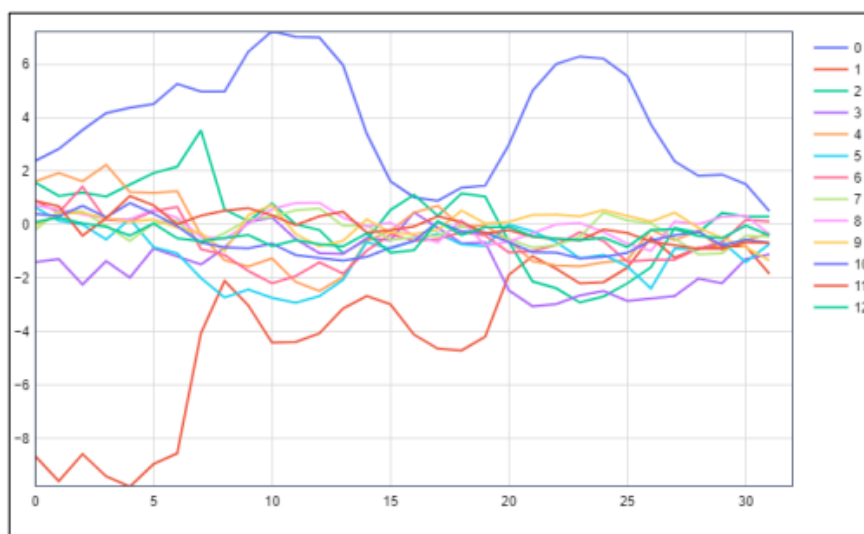


Рисунок 2.29 – Приклад датасету SpokenArabicDigits

StandWalkJump. Датасет визначення рухів людини за допомогою електрокардіограми. Дані є спектрограми, записані при певних рухах випробуваного. Приклад датасету показано на рис 2.30.

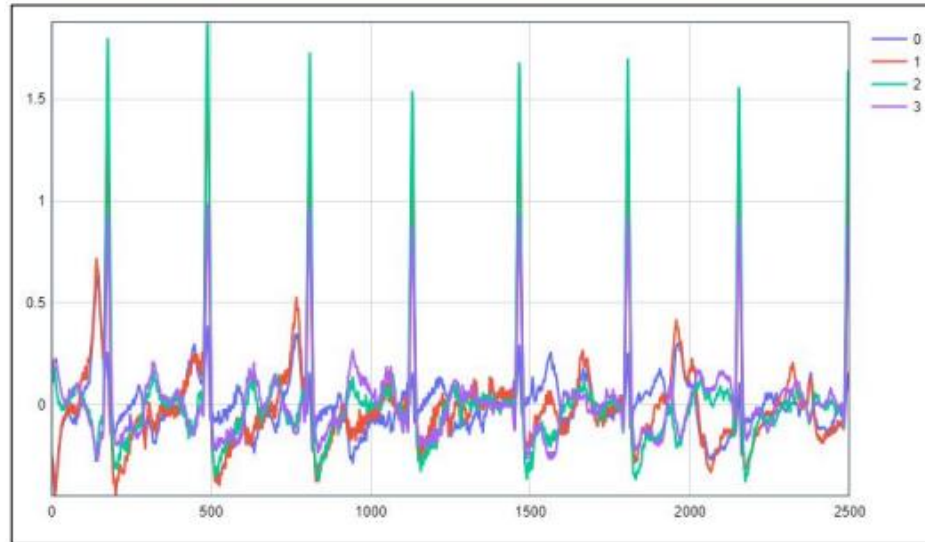


Рисунок 2.30 – Приклад датасету StandWalkJump

InsectWingBeat. Датасет з визначення виду москітів по звуку, що видається. Дані збережені у вигляді спектрограм та характеризують собою потужність звуку. Кожен клас характеризує одну з особливостей виду москітів. Приклад датасета подано на рис. 2.31

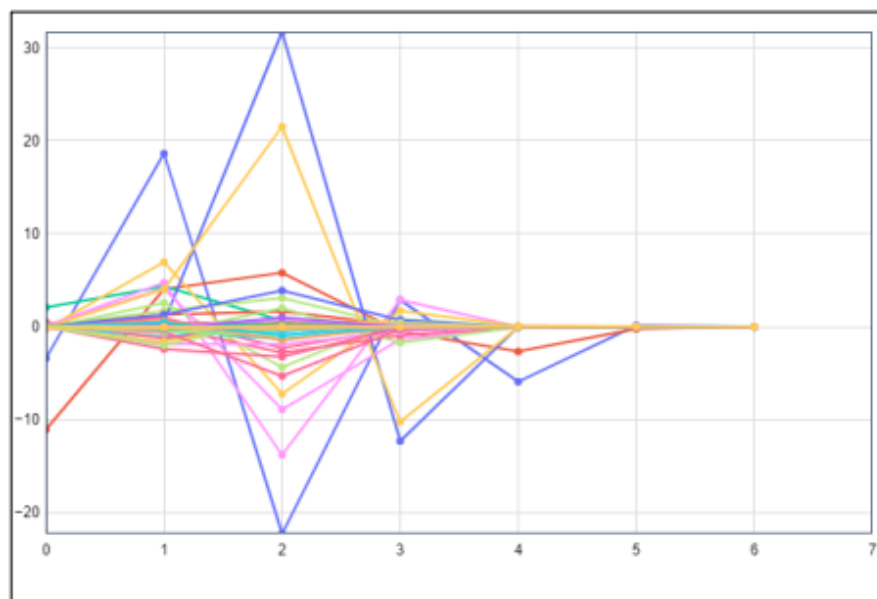


Рисунок 2.31 – Приклад датасету InsectWingBeat

UWaveGestureLibrary. Датасет за визначенням жестів на різних девайсах за допомогою тривісних акселерометрів. Дані були зібрані з різних електронних пристроїв із функцією виконання команд за допомогою жестів пристрою. Приклад датасету наведено на рис. 2.32

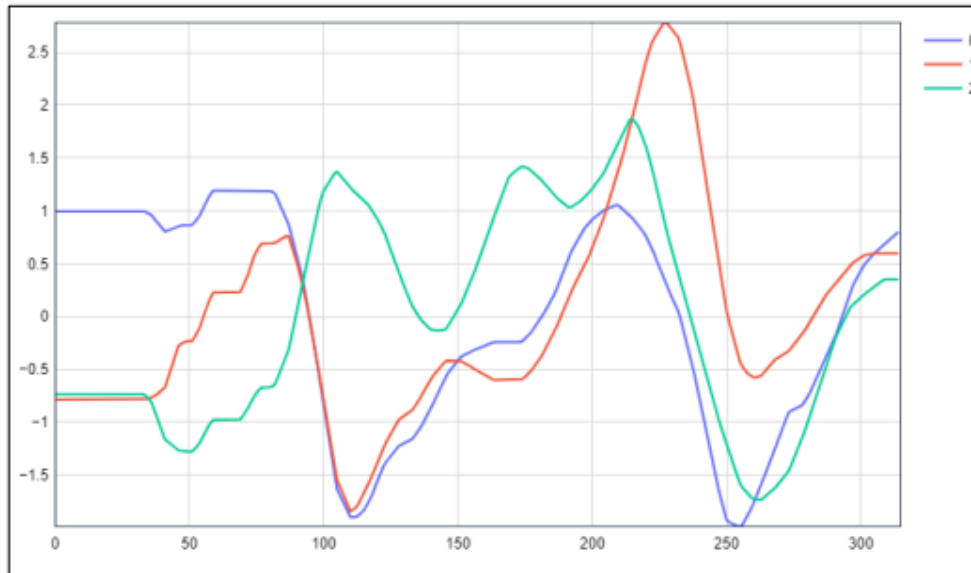


Рисунок 2.32 – Приклад датасету UWaveGestureLibrary

2.3 Висновки до другого розділу

У цьому розділі було розроблено архітектуру моделі трансформера, котра дає змогу враховувати не тільки взаємозв'язки між локальними областями серед кожного із вимірів, але й взаємозв'язки між усіма підпоследовностями властиво вздовж кожного виміру.

Для проведення оцінки запропонованої архітектури задачі класифікації БЧР було визначено критерії відбору даних. Під встановлені критерії потрібно використати набір даних UEA. Для якіснішого висвітлення результатів тестування побудованої архітектури було проаналізовано кожен датасет з набору UEA.

3 ПРАКТИЧНА ЧАСТИНА

3.1 Інструмент автоматичного підбору гіперпараметрів

Для автоматичного підбору найкращих параметрів передобробки та гіперпараметрів розробленої моделі класифікації БЧР розроблено модуль автоматичного перебору різних параметрів [38].

Вхідні дані навчання та тестування приймаються у модуль у форматі датасетів UEA (ARFF).

3.1.1 Алгоритм роботи інструменту

Цей модуль реалізований мовою програмування Python [39] та виконує послідовний перебір параметрів для досягнення найкращих результатів. Основна ідея полягає у переборі кожного гіперпараметра та збереження значення, за допомогою якого було отримано найбільшу метрику на контрольній вибірці. Як метрика оцінки використана частка правильних відповідей серед усіх прикладів.

Початкові значення гіперпараметрів та його діапазони перебору представляються у файлі конфігурації. Цей файл приймається у форматі JSON . Також у файлі конфігурації представлені обмеження гіперпараметрів для ефективного підбору при вирішенні завдань з більшими обсягами даних [38].

Далі представлений алгоритм виконання автоматичного перебору параметрів передобробки та гіперпараметрів моделі.

1. Ініціалізація початкових параметрів. На даному етапі відбувається встановлення початкових значень параметрів та гіперпараметрів, які будуть використовуватись у подальшому процесі перебору. Ініціалізація початкових параметрів потрібна для того, щоб задати відправні точки для подальшого перебору.

2. Перебір параметрів нормалізації. На цьому етапі перебираються

різні методи нормалізації даних. Використовується нормалізація за мінімумом та максимумом. Проводиться перебір таких варіантів як нормування кожного виміру окремо, нормування щодо глобального мінімуму та максимуму, не використання нормалізації. Нормалізація даних допомагає покращити збіжність алгоритмів навчання та збільшити точність моделі.

3. Використання перетворення Фур'є. На цьому етапі розглядається застосування перетворення Фур'є до даних, що дозволяє перейти від часової до частотної області. Перетворення Фур'є може виявити важливі частотні компоненти даних, які можуть бути корисні для поліпшення якості моделі.

4. Перебір варіантів об'єднання та вибору вимірювань. Перебираються різні стратегії об'єднання та вибору вимірів, такі як об'єднання всіх вимірів в один масив, вибірка найбільш значимих вимірів та інші підходи. Правильний вибір та поєднання вимірів може значно вплинути на продуктивність моделі. Це допомагає зменшити розмірність даних та покращити ефективність навчання.

5. Перебір розміру патчів. На цьому етапі перебираються різні розміри патчів, які будуть використовуватися при розбитті даних на дрібніші частини. Розмір патчів може впливати на здатність моделі захоплювати локальні патерни даних. Правильний вибір розміру патчів може покращити навчання моделі та її здатність до узагальнення. Розміри для перебору залежать від довжини вхідної послідовності та є дільниками довжини.

6. Перебір кроку навчання. Перебираються різні значення кроку навчання, який визначає, наскільки модель оновлює свої ваги при кожному кроці навчання. Крок навчання є одним із ключових гіперпараметрів, який впливає на збіжність та стабільність процесу навчання. Правильний крок навчання допомагає уникнути як застрягання в локальних мінімумах, так і перескакування через оптимальні значення. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 0.00001, 0.0001, 0.001, 0.01.

7. Перебір кількості нейронів у повнозв'язкових шарах енкодерів. Правильний вибір цього параметра допомагає досягти балансу між складністю

моделі та її здатністю до узагальнення. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 516, 1024, 2048.

8. Перебір частки занулення нейронів (Dropout). Перебираються різні частки занулення нейронів в енкодер шарах моделі. Даний гіперпараметр дозволяє зменшити перенавчання та підвищити стійкість моделі. Можливі значення для перебору можуть встановлюватись користувачем. У роботі були використані значення 0.1, 0.25, 0.5, 0.75.

9. Перебір кількості енкодерів. Перебираються різні кількості енкодерів, що використовуються в моделі. Кількість енкодерів впливає на глибину моделі та її здатність захоплювати складні патерни у даних. Перебір дозволяє знайти оптимальну кількість енкодерів для вирішення задачі. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 0, 1, 2, 3.

10. Використання шару ембедингів. Розглядається включення чи виключення шару ембедингів до архітектури моделі. Ембединги можуть поліпшити представлення даних моделі і це може підвищити точність моделі.

11. Перебір кількості голів уваги. Перебираються різні кількості голів уваги моделі. Кількість голів уваги визначає, скільки різних аспектів інформації модель може одночасно враховувати. Це може значно покращити здатність моделі до навчання та її продуктивність на складних завданнях. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

Перебір вищевикладених гіперпараметрів у представленій послідовності дозволяє досягти найкращих результатів для вирішення задачі класифікації БЧР за допомогою розробленої моделі трансформера [38].

3.1.2 Інтерфейс інструменту

Для зручності використання підбору гіперпараметрів було написано

графічний інтерфейс за допомогою використання бібліотеки GRADIO [40]. Розроблений інтерфейс дозволяє зробити розрахунки у трьох режимах.

1. Train and Test. Даний режим дозволяє підібрати найкращі гіперпараметри моделі та передобробки даних, зберегти знайдені параметри та навчену модель, а також вивести метрику на тестовій вибірці.

Необхідні вхідні дані:

- набір даних для навчання;
- набір даних для тестування;
- конфігураційний файл;
- шлях у файловій системі для збереження результатів.

Вихідні дані:

- навчена модель у форматі PyTorch [41];
- найкращі гіперпараметри моделі та передобробки даних;
- отримана метрика на тестовій вибірці.

Вікно режиму показано на рис. 3.1.

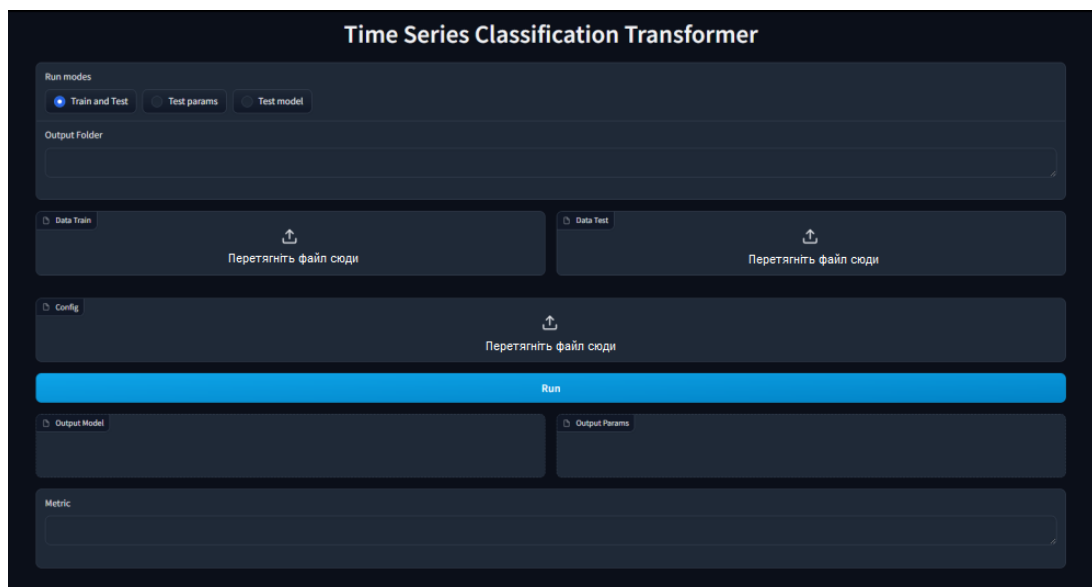


Рисунок 3.1 – Вікно режиму Train and Test

2. Test params. Даний режим дозволяє протестувати параметри попередньої обробки та гіперпараметри моделі з файлу XLSX. Режим навчає модель з отриманими гіперпараметрами та видає метрику на тестовій вибірці та

збережену модель.

Необхідні вхідні дані:

- гіперпараметри моделі та передоброби даних;
- набір даних для навчання;
- набір даних для тестування;
- конфігураційний файл;
- шлях у файловій системі для збереження результатів.

Вихідні дані:

- навчена модель у форматі PyTorch [41];
- отримана метрика на вибірці.

Вікно режиму показано на рис. 3.2.

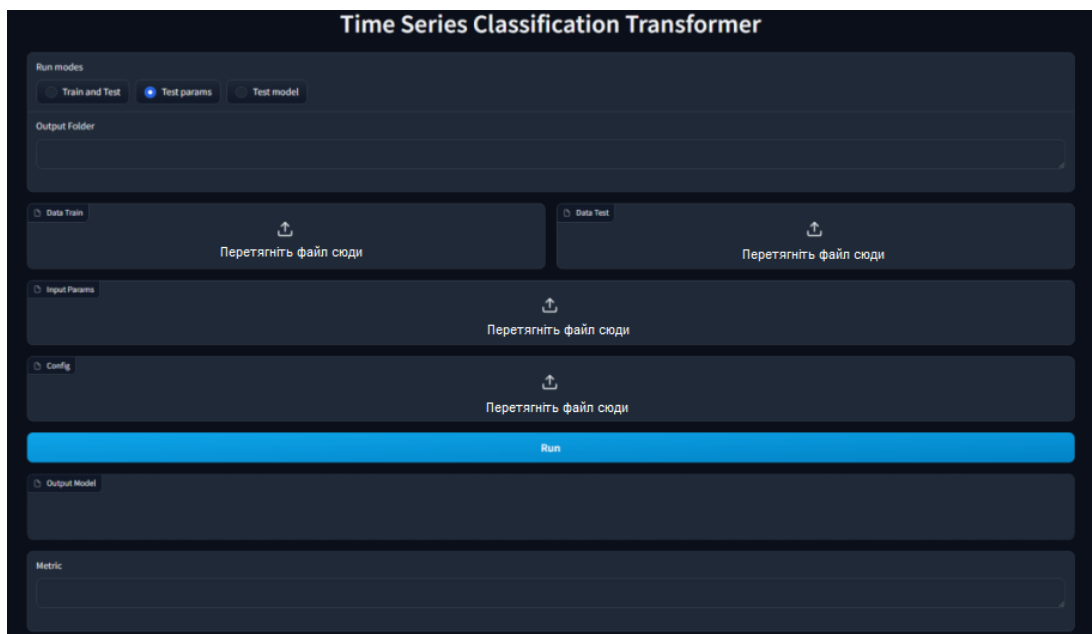


Рисунок 3.2 – Вікно режиму Test params

3. Test model. Цей режим дозволяє протестувати модель, яка завантажується у форматі PyTorch. Режим використовує збережені в моделі ваги та видає результат на тестовій вибірці.

Необхідні вхідні дані:

- гіперпараметри моделі та передоброби даних;
- набір даних для тестування;

- конфігураційний файл;
- файл моделі класифікації;
- шлях у файлової системі для збереження результатів.

Вихідні дані:

- отримана метрика на вибірці.

Вікно режиму відображено на рис. 3.3.

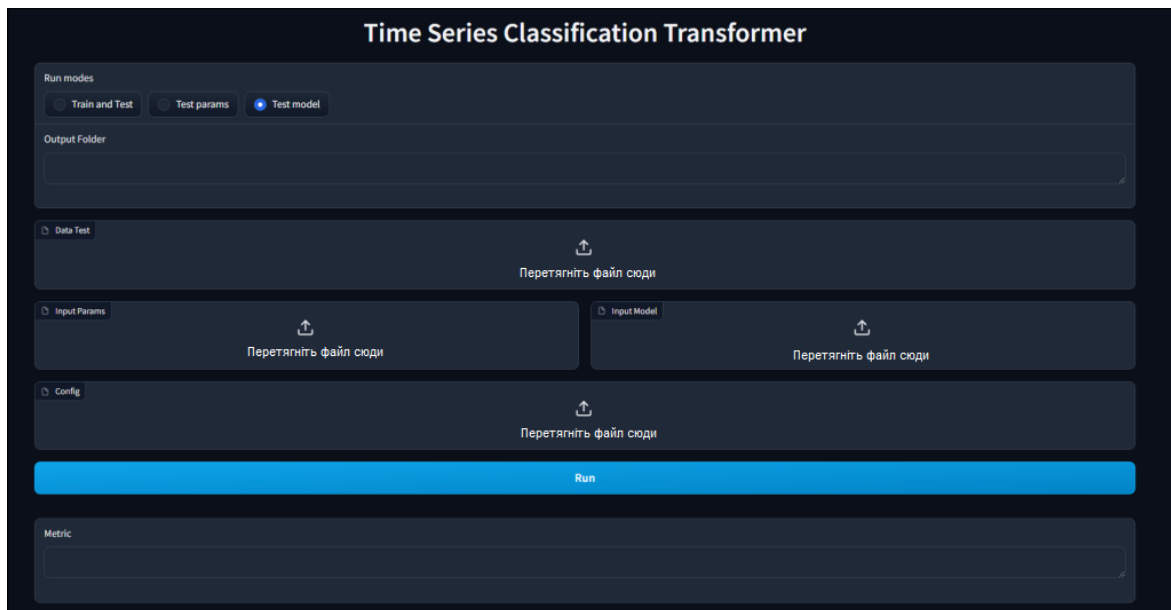


Рисунок 3.3 – Вікно режиму Test model

3.1.3 Алгоритм використання інструменту

На першому етапі необхідно виконати вибір режиму. Вікно для виконання цієї дії наведено на рис. 3.4.

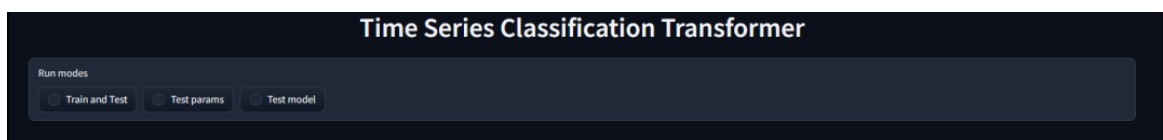


Рисунок 3.4 – Вікно вибору режиму

На другому етапі потрібно завантажити вхідних даних (рис. 3.5). Залежно від режиму змінюються вхідні дані.

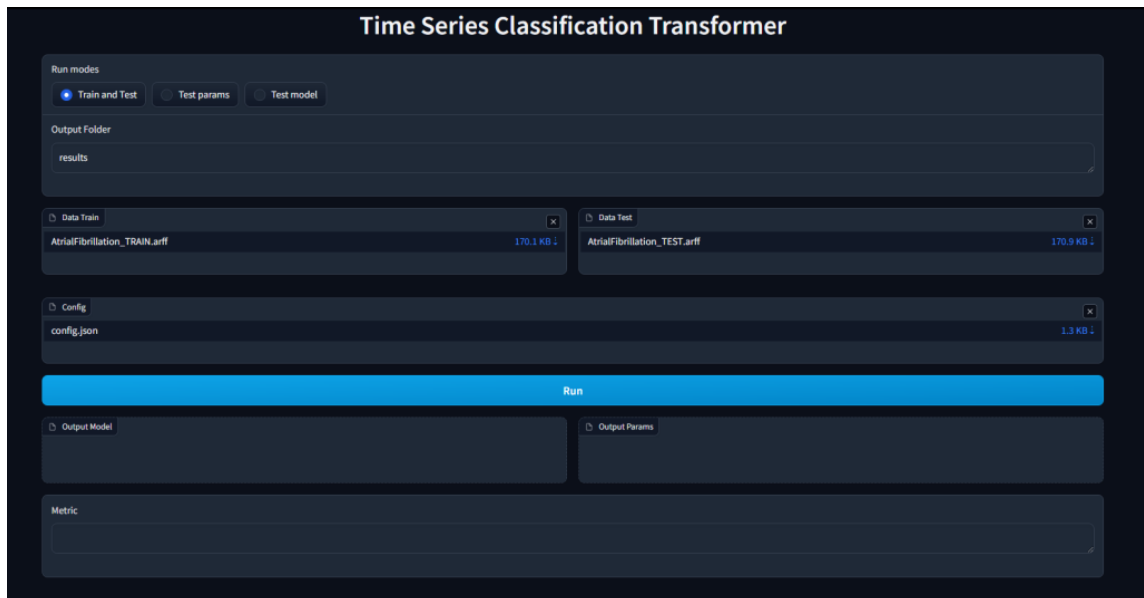


Рисунок 3.5 – Вікно завантаження даних

Потім треба виконати розрахунок та отримати вихідні файли та метрики якості. Залежно від режиму можуть бути різні вихідні дані. Вікно виведення результатів показано на рис. 3.6.

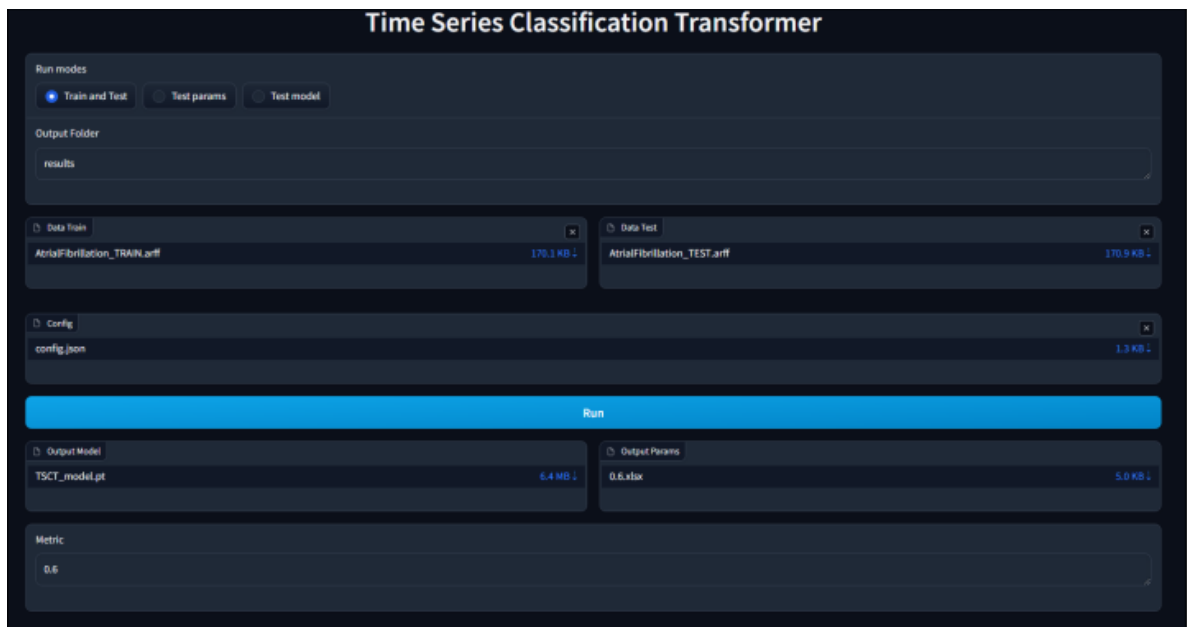


Рисунок 3.6 Вікно виведення результатів

3.2 Тестування розробленої моделі

Для оцінки якості розробленої моделі класифікації БЧР було проведено

підбір гіперпараметрів, навчання та тестування для всіх датасетів набору UEA [26]. Авторами датасетів було зроблено поділ на навчальну та тестову вибірки, який використовувався надалі при тестуванні розробленої моделі.

Тестування проводилося на локальному комп'ютері, характеристики якого представлені в таблиці 3.1.

Таблиця 3.1– Обчислювальні характеристики пристрою тестування

Характеристика	Найменування
Процесор	AMD Ryzen 7 4800H with Radeon Graphics 2.90 GHz
Відеокарта	NVIDIA GeForce RTX 3050 Laptop GPU
Обсяг оперативної пам'яті	16 Гб
Обсяг графічної пам'яті	4 Гб
ОС	Windows 11

Для підбору гіперпараметрів розробленої моделі було виділено контрольну вибірку, що становить 10 % від навчальної вибірки. Для ефективного підбору параметрів попередньої обробки та гіперпараметрів моделі був використаний інструмент, описаний у підрозділі 2.2.

Тестування проводилося за таким алгоритмом.

- завантаження навчальної та тестової вибірок;
- виділення з навчальної вибірки контрольної частини;
- запуск підбору гіперпараметрів;
- запуск моделі з найкращими гіперпараметрами на тестовій вибірці;
- збереження результатів.

Ця стратегія тестування забезпечує незалежну оцінку ефективності та точності розробленої моделі, що дозволяє не лише підтвердити її працездатність, а й виявити ключові недоліки в архітектурі.

Як метрика для оцінки алгоритмів, була обрана частка правильних відповідей серед усіх прикладів (ассурасу). Дана метрика ідеально підходить для

оцінки якості моделі на наборі даних, що тестується, так як всі класи у вибірках розподілені рівномірно. І оскільки немає дисбалансу класів асигасу показує оцінку без усунення тому чи іншого класу.

Для оцінки відносної помилки було проведено порівняння із SOTA моделлю, якою на даний момент є DyFormer [32]. У таблиці 3.2 представлені результати розробленої моделі та найкращої у науковому співтоваристві на тестовій вибірці. Потовщенням виділені датасети в яких розроблена модель показала однаковий або більший результат порівняно з SOTA моделлю. Також у табл. 3.2 наведено середні точності по обох моделях.

Таблиця 3.2 - Точність за всіма прикладами SOTA та розробленої моделі

Датасет	DyFormer	Розроблена модель
ArticularyWordRecognition	0,997	0,9933
AtrialFibrillation	0,267	0,6
BasicMotions	1	1
CharacterTrajectories	0,999	0,9854
Cricket	1	0,9861
DuckDuckGeese	0,788	0,82
EigenWorms	0,901	0,92
Epilepsy	1	1
Ering	0,97	0,9704
EthanolConcentration	0,397	0,3802
FaceDetection	0,732	0,6833
FingerMovements	0,65	0,69
HandMovementDirection	0,459	0,7162
HandWriting	0,593	0,3165
HeartBeat	0,816	0,7805
JapaneseVowels	0,989	0,9703
Libras	0,944	0,9111
LSST	0,758	0,6176

1	2	3
MotorImagery	0,64	0,68
NATOPS	0,933	0,9611
PEMS-SF	0,952	0,8382
PenDigits	0,997	0,9626
Phoneme	0,309	0,2183
RacketSports	0,954	0,9013
SelfRegulationSCP1	0,944	0,9386
SelfRegulationSCP2	0,628	0,6333
SpokenArabicDigits	0,996	0,9909
StandWalkJump	0,433	0,7333
UWaveGestureLibrary	0,962	0,9062
InsectWingBeat	0,711	0,7232
Середня точність	0,7906	0,7943

Аналізуючи отримані результати, можна зробити такі висновки. По 13 з 30 датасетів розроблена модель отримала більш високу точність, ніж найкраща модель у науковому співтоваристві. Поліпшення середньої точності становило 0,37 %. Поліпшення простежується за тими датасетами, у яких ключову роль класифікації відіграють взаємозв'язку між вимірами.

3.3 Вплив особливостей розробленої моделі

3.3.1 Вплив розміру патчу

Розмір патча може істотно впливати на результат, оскільки основна ідея архітектури полягає в пошуку закономірностей між патчами. Для оцінки впливу цього параметра було складено таблиці 3.3. в якій представлені розміри патчів,

що показали найкращу метрику по кожному з датасетів тестової вибірки. Для проведення аналізу дані у табл. 3.3 відсортовано за довжинами послідовностей

Таблиця 3.3 - Підібрані розмірності патчів та послідовностей по кожному з датасетів у тестовій вибірці

Датасет	Розмір патча	Довжина послідовностей	% розміру патча від довжини
PenDigits	4	8	50,00
JapaneseVowels	29	29	100,00
InsectWingBeat	11	30	36,67
RacketSports	6	30	20,00
LSST	6	36	16,67
Libras	9	45	20,00
FingerMovements	10	50	20,00
NATOPS	17	51	33,33
FaceDetection	31	62	50,00
Ering	13	65	20,00
SpokenArabicDigits	31	93	33,33
BasicMotions	10	100	10,00
ArticulatoryWordRecognition	12	144	8,33
PEMS-SF	12	144	8,33
HandWriting	19	152	12,50
CharacterTrajectories	14	182	7,69
Epilepsy	103	206	50,00
PhonemeSpectra	31	217	14,29
DuckDuckGeese	18	270	6,67
UWaveGestureLibrary	21	315	6,67
HandMovementDirection	20	400	5,00
HeartBeat	27	405	6,67
AtrialFibrillation	32	640	5,00

1	2	3	4
SelfRegulationSCP1	32	896	3,57
SelfRegulationSCP2	36	1152	3,13
Cricket	57	1197	4,76
EthanolConcentration	103	1751	5,88
StandWalkJump	50	2500	2,00
MotorImagery	60	3000	2,00
EigenWorms	281	17984	1,56

Аналізуючи наведену таблицю, можна дійти висновку, що з збільшенням розміру послідовностей необхідно зменшувати розмір патчів. Це пояснюється тим, що менші патчі дозволяють моделі більш точно захоплювати і обробляти локальні залежності даних, що особливо важливо під час роботи з довгими послідовностями. Більші патчі можуть втрачати важливі деталі та контекст, що призводить до погіршення якості пророцтв.

Також для визначення впливу розміру патча на результат було проведено експеримент на датасеті StandWalkJump, оскільки по ньому виявлено суттєве поліпшення результатів (див. табл. 3.2). У ході експерименту проводилося навчання з різними розмірами патчів та записувався результат навчання. Інші гіперпараметри моделі були постійними. Результат експерименту показано на рис. 3.7. Для наочності вісь абсцис наведена у логарифмічній шкалі.

Таку ж залежність зміни результату від розміру патчу можна побачити на більшості датасетів. Ця залежність підтверджує, що розроблена архітектура була побудована правильно. Так як при малому та великому розмірі патча спостерігається низькі результати, що означає, що модель не може коректно знайти патерни за таких розмірностей.

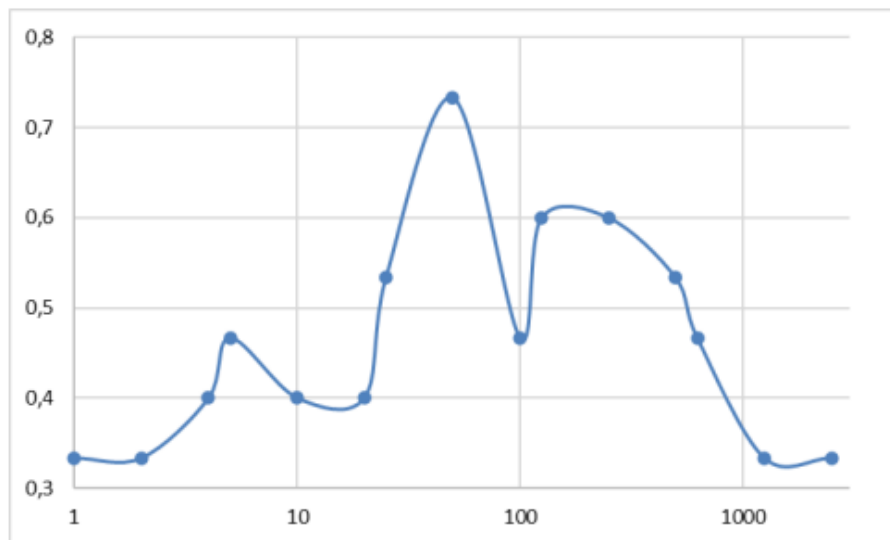


Рисунок 3.7 – Зміна точності залежно від розміру патча

3.3.2 Вплив кількості енкодерів

Енкодери у розробленій архітектурі є ключовою ланкою, що виконує основне завдання – пошук закономірностей. У розробленій архітектурі є два типи енкодерів. Перший шукає закономірностей між вимірами, а другий шукає взаємозв'язку між патчами вздовж кожного виміру. Кількість кожного типу енкодера є гіперпараметром і може змінюватись. Для оцінки впливу кількості енкодерів на результат була складена таблиці 3.4, в яких відображені найкращі варіанти даного параметра по кожному з датасетів у тестовій вибірці.

Таблиця 3.4 – Кількість підібраних енкодерів по кожному з датасетів тестової вибірки

Датасет	Кількість енкодерів 1	Кількість енкодерів 2
ArticularyWordRecognition	1	1
AtrialFibrillation	1	1
BasicMotions	1	1
CharacterTrajectories	1	1
Cricket	1	1
DuckDuckGeese	1	1
EigenWorms	1	0

1	2	3
Epilepsy	1	1
Ering	1	1
EthanolConcentration	1	1
FaceDetection	1	1
FingerMovements	1	1
HandMovementDirection	1	1
HandWriting	3	1
HeartBeat	3	2
InsectWingBeat	1	1
JapaneseVowels	1	1
Libras	3	3
LSS1	1	2
MotorImagery	1	1
NATOPS	1	3
PEMS-SF	0	1
PenDigits	3	2
PhonemeSpectra	3	2
RacketSports	2	0
SelfRegulationSCP1	1	1
SelfRegulationSCP2	1	1
SpokenArabicDigits	1	2
StandWalkJump	2	2
UWaveGestureLibrary	0	3

Аналізуючи табл. 3.4, можна зробити такі висновки.

Збільшення енкодерів типу 1 сприяє більш глибокому виявленню закономірностей між вимірами, а збільшення енкодерів типу 2 сприяє виявленню залежностей вздовж кожного виміру часових рядів.

У переважній більшості датасетів найкращою кількістю енкодерів кожного типу є 1. Але є також є датасети в яких той чи інший тип енкодера

переважає. Наприклад, датасет HandWriting. У ньому кількість енкодерів типу 1 - 3, а кількість енкодерів типу 2 - 1. У даній мережі вирішувалося завдання визначення написаної людиною букви за допомогою акселерометрів, що записують рух руки при написанні букви по трьох координатних осях. Аналізуючи отриманий результат моделлю, з'ясовується, що модель «зрозуміла» необхідність урахування зміни вимірів щодо один одного, що насправді характеризує тривимірний простір. І це виявилось важливішим за облік зміни в часі, що теж не виключає модель використовуючи енкодер типу 2. Зворотний результат можна побачити в датасеті UWaveGestureLibrary. Тут кількість енкодерів типу 1 - 0, а кількість енкодерів типу 2 - 3. У даній мережі також визначається клас по тривісному акселерометру, а саме необхідно визначити форму жесту. У цьому завданні модель визначила, що форму жесту характеризує зміна координат за часом, а чи не зміни вимірів щодо один одного..

3.4 Висновки до третього розділу

Третій розділ присвячений практичній реалізації програмного інструменту для підбору гіперпараметрів у автоматичному режимі. Описано алгоритми його роботи та використання, а також інтерфейс інструменту.

Проведено тестування розробленої моделі, досліджено вплив розміру патчу та кількості енкодерів на модель.

Аналізуючи отримані результати на схожих датасетах, можна дійти висновку, що для виявлення складних закономірностей (класифікація букв) необхідно орієнтуватися зміни вимірів один щодо одного, а виявлення простіших закономірностей (класифікація жестів) необхідно орієнтуватися зміну вздовж кожного часового ряду, що дозволяє досягти розроблена модель.

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Закордонний досвід організації охорони праці в ІТ-компаніях

Стан справ з охороною праці у світі стає все більш актуальною проблемою як для профспілок, так і для міждержавних структур, насамперед Міжнародної організації праці (МОП). МОП розглядає цю тему як частину своєї Програми гідної праці. Підвищена увага до проблем безпеки праці пояснюється в першу чергу тим, що з кожним роком, незважаючи на заходи, що вживаються, у різних країнах зростає рівень виробничого травматизму, у тому числі зі смертельними наслідками, і кількість профзахворювань.

На перший погляд, робота за комп'ютером при дослідженні моделей та побудові системи ПФПК здається безпечною, але саме легковажність до неї може призвести до певних проблем у здоров'ї людини. Професія програміста та інших фахівців ІТ-технологій пов'язана з колосальним розумовим напруженням. Розробники – настільки захоплені люди, що навіть відволікаючись від роботи над проектом, продовжують думати про роботу. Нерідко відпочинком вони вважають паралельну заміну основної діяльності, наприклад, читання профільної літератури, верстку сайтів, вивчення нових мов програмування. Однак мозок не може до безкінечності приймати виключно корисну інформацію, яку розробник прагне направляти в русло особистісного та професійного зростання [28].

Зарубіжний досвід охорони праці при використанні новітніх інформаційних технологій та сучасного комп'ютерного обладнання передбачає з метою попередження наслідків монотонної праці, підвищення рівня рухової активності і покращення розумової працездатності фахівців ІТ-індустрії під час технологічних перерв участь у спеціальних облаштованих приміщеннях необхідним спортивним інвентарем та різними тренажерами відповідних фізичних вправ, індивідуальних тренінгових завдань відповідно до віку, статі та категорії зорової роботи. Такий підхід дозволяє зняти надлишкове

психофізіологічне перевантаження, підвищити працездатність центральної нервової системи, попередити перевтому зорового аналізатора. Показана ефективність проведення різноманітних за своєю спрямованістю вправ робітників цієї галузі (приблизно на 5-30%) [28].

Зараз багато ІТ-компаній обладнують свої офіси кімнатами відпочинку та лаунж-зонами, які забезпечують психофізіологічне розвантаження працівників. Адже окремим робочим столом з ноутбуком вже давно нікого не здивуєш. Тому, бажаючи підвищити продуктивність працівників, міжнародні компанії змагаються, перетворюючи нудні одноманітні офіси в креативні простори, де нові ідеї народжуються без титанічних зусиль. Наприклад, винахідники компанії Inventionland трудяться в казкових декораціях. Тут і гігантський гоночний трек, і пряниковий будиночок, і дуже реалістичний піратський корабель, на палубі якого розташувалися комп'ютерні столи, ніжками яких служать винні бочки. Робочі місця співробітників компанії Google в Цюриху нагадують гігантські вулики, а офіс шведського інтернет-провайдера Bahnhof розташувався в бомбосховищі часів холодної війни і походить на підземний притулок землян після глобальної катастрофи. А щоб співробітників не тягнуло додому, роботодавці створюють і можливість релаксувати, не відходячи від робочого місця, обладнавши басейни, ігрові кімнати та спортзали [28].

Корпорація Google піклується і про санітарно-гігієнічні умови праці та використовує систему очищення повітря, яка видаляє з атмосфери всі токсини та важкі домішки [29]. Також незвичайними на території офісу є спеціальні ізолюючі світло й звук капсули, де втомлений співробітник може відпочити, так само як і в спеціалізованих кімнатах у зоні відпочинку з приглушеним світлом та заспокійливою музикою, масажних кабінетах, ігрових кімнатах. Культивуваці активного відпочинку та спорту в Google відводиться особлива увага, оскільки одні з найпоширеніших хвороб програмістів пов'язані зі спиною. Саме тому в головному офісі міститься спортзал, в якому можуть одночасно займатися велика частка співробітників, до того ж на території є плавальний басейн з черговим рятувальником, тому, фактично, людина може працювати прямо

звідти (пам'ятаючи про дотримання правил техніки безпеки) [30]. Для переміщення по великій території Гуглмістечка «гуглівці» (так себе називають співробітники цієї компанії) користуються самокатами, які мають моторчики. Звичайно, усі ці зручності безкоштовні.

Намагається не відставати від Google і корпорація Facebook. Так, дана фірма пропонує своїм співробітникам чудові умови праці: якісне різноманітне харчування, зручне апаратне забезпечення, важливою складовою якого є широкі монітори ПК з високою роздільною здатністю [31]. Особливістю Facebook є те, що співробітники не мають ані своїх кабінетів, ані офісних комірок – усі сидять разом, оскільки це стимулює комунікацію та обмін ідеями, підвищує продуктивність роботи. Загалом у компанії царює атмосфера позитиву. Усі співробітники привітні та посміхаються. На стінах можна побачити карикатури на топменеджерів, а робітники одягають футболки зі смішними написами [31]. Співробітники тут, на відміну від Google, пересуваються по офісу не на самокатах, а на скейтах. У Facebook також наявна велика кількість тематичних та ігрових зон, де кожен може розслабитися та відпочити, і це все також абсолютно безкоштовно.

Поява та впровадження нових інформаційно-комунікаційних технологій зумовлює необхідність подальшого вдосконалення охорони праці фахівців ІТ-індустрії. З метою належного правового забезпечення необхідно розширити та доповнити перелік основних професій комп'ютерної галузі у національному класифікаторі [32], а також підготувати відповідний випуск у кваліфікаційному довіднику посад фахівців ІТ-індустрії, що сприятиме вирішенню питань їх соціального захисту, пенсійного забезпечення, атестації робочих місць основних професій за умовами праці на предмет подальших певних видів пільг та компенсацій за важкі шкідливі і небезпечні умови праці. Важливим напрямом стосовно визначення професійної придатності фахівців з ІТ є проведення психофізіологічної експертизи відповідно до 5 статті [33].

ІТ-фахівці, як і будь-які інші працівники, повинні проходити навчання і перевірку знань з охорони праці або в навчальному центрі, або в самій

організації. Якщо в ній є комісія з перевірки знань з охорони праці, атестованих в спеціалізованому навчальному центрі. Навчання охорони праці в організації проводять по самостійно розробленими програмами. Їх складають, спираючись на типові програми, а також з огляду на особливості галузі, в якій працює організація.

Робота з комп'ютерами нового покоління характеризується певним психофізіологічними перенавантаженнями, втому зорового аналізатора, гіпокінезією, відсутність диференційованих норм праці при роботі з новою комп'ютерною технікою в залежності від віку, статі, категорії зорової роботи, режимів праці і відпочинку (протягом робочого дня, тижня, щорічного режиму відпусток). Все це потребує розробки нових нормативно-правових актів з регламентації праці та відпочинку фахівців ІТ-індустрії і стандартів підприємств, центрів комп'ютерної техніки, центрів інформаційних технологій, сучасних комп'ютерних класів. Особлива роль з точки зору збереження та відновлення здоров'я працюючих в комп'ютерній галузі належить попереднім та періодичним наглядам з подальшої психофізіологічної експертизи і встановленням професійної придатності при роботі з комп'ютерами нового покоління, який супроводжується виникненням певних факторів професійного ризику електротравматизму при їх ремонті та обслуговуванні. В цьому зв'язку необхідне запровадження експертизи на предмет безпечної експлуатації ПЕОМ, тобто офіційне підтвердження фактичних параметрів електробезпеки, їх відповідності вимогам нормативної документації фахівців, які проводять таку експертизу повинні пройти навчання і перевірку знань відповідно до вимог [34]. За результатами експертизи повинні прийматися рішення про відповідність ПЕОМ нормам безпеки, терміни чергової експертизи, оформлюються протоколи вимірювань і випробувань, проведені у разі потреби розрахунки та експертний висновок. Для підвищення розумової працездатності то зорової роботи повинна здійснюватися ергономічна оптимізація в рамках системи «оператор-термінал», яка сприятиме результативній фізичній та інтелектуальній працездатності і відновленню психосоматичного здоров'я фахівців ІТ-індустрії.

Заслуговує на увагу зарубіжний досвід створення у приміщеннях та в зоні їх розміщення на територіях підприємств спеціальних візуальних комфортних умов та забезпечення вимог виробничої естетики, дотримання норм рівнів виробничого шуму та акустичної тиші за межами офісу. Також дуже важливим є використання в офісних приміщеннях та кабінетах психофізіологічного розвантаження функціональної музики, яка сприяє попередженню перевтоми і підтриманню необхідного рівня розумової працездатності фахівців комп'ютерної галузі. В цьому напрямі заслуговує на увагу створення при великих центрах інформаційних технологій кімнат (кабінетів) психофізіологічного розвантаження працівників галузі (на 5 місць) [29].

Всі наведені заходи щодо вдосконалення охорони праці фахівців ІТ-індустрії повинні контролюватися службою охорони праці та комісією з охорони праці підприємства. Особливе значення у соціальному захисті цієї категорії працівників належить прийняття комплексного договору, який може забезпечити фахівців додатковими пільгами та компенсаціями.

Таким чином, використання новітніх технологій вимагає від фахівців ІТ-індустрії дотримання певних правил та вимог з точки зору безпеки праці, її нормування з урахуванням віку працюючих та загального інформаційного навантаження, розробки та впровадження індивідуальних, щотижневих та щорічних режимів праці та відпочинку, які сприятимуть профілактиці перевтомлення і підвищенню розумової працездатності працюючих. Особливу роль в цьому напрямі повинні відігравати ергономічні заходи стосовно створення робочих місць, оптимізації взаємодії людини в рамках системи «оператор-термінал».

4.2 Оцінка дії електромагнітного імпульсу на елементи комп'ютерної системи

У воєнний час при застосуванні ядерної зброї проти України на електронно-обчислювальне обладнання в першу чергу буде впливати електромагнітний

імпульс (ЕМІ) ядерного вибуху у вигляді короткого імпульсу, який вражає головним чином електричну та електронну апаратуру [35].

ЕМІ виникають в основному в результаті взаємодії гамма-випромінювання з атомами навколишнього середовища. На утворення ЕМІ йде невелика кількість ядерної енергії, але він здатен викликати високі імпульси струмів та напруг в кабелях повітряних і підземних ліній зв'язку, сигналізації, управління, електропередачі, в антенах радіостанцій. Вплив ЕМІ може привести до згорання чутливих електронних та електричних елементів, зв'язаних з великими антенами чи відкритими дротами, а також до порушень в обчислювальних пристроях.

Особливістю ЕМІ, як вражаючого фактору є його здатність розповсюджуватись на десятки і сотні кілометрів в оточуючому середовищі. Тому ЕМІ може вплинути своєю дією на об'єкти, там де вибухова хвиля, світлове випромінювання, проникаюча радіація втрачають своє значення, як вражаючі фактори. При наземних та низьких повітряних вибухах в лініях зв'язку та електрозабезпечення виникають напруги, які можуть викликати пробій ізоляції провідників та кабелів відносно землі, пробій ізоляції елементів приладів підключених до повітряних і підземних ліній. Степінь враження залежить від наведеного імпульсу напруги чи струму і також електричної міцності обладнання.

Найбільш піддані впливу ЕМІ системи зв'язку, сигналізації, управління. Використані в цих системах кабелі та апаратура мають обмежену електричну міцність не більше 10кВ імпульсної напруги, тоді як наведені імпульси напруги від ЕМІ можуть перевищувати ці значення. Найбільш піддана впливу ЕМІ апаратура виконана на напівпровідниках та інтегральних схемах, працюючих на малих струмах і напругах, і значить відчутних до впливу зовнішніх електричних і магнітних кіл, в тому числі і елементи програмного засобу для управління процесом міграції віртуальних машин в обчислювальній хмарі. ЕМІ пробиває ізоляцію, спалює елементи електричних схем радіоапаратури, викликає коротке замикання в радіопристроях, іонізацію діелектриків, змінює або повністю стирає магнітний запис. Встановлено, що при дії ЕМІ на апаратуру найбільша напруга

наводиться на вході. В транзисторах відбувається така залежність: чим більший коефіцієнт підсилення транзистора, тим менша його електрична міцність.

ЕМІ пошкоджує також резистори, викликає іскріння в їх міжконтактних з'єднаннях і деяких областях провідної поверхні. Найбільшу небезпеку ЕМІ представляє для апаратури, яка встановлена в особливо міцних спорудах, які витримують великі тиски ударної хвилі. В цих спорудах апаратура не виходить з ладу від механічних пошкоджень, але ЕМІ може вивести з ладу всю незахищену апаратуру системи зв'язку, сигналізації і керування. Найбільших значень досягають напруги, які наводяться між кабелем і землею.

Розглянемо можливі шляхи рішення задачі захисту від ЕМІ сервісу для адміністрування і обліку роботи автомобільної пар. Ідеальним захистом від ЕМІ виявилось б повне укриття приміщення, в якому розміщена радіоелектронна апаратура, металевим екраном. Водночас зрозуміло, що практично забезпечити такий захист у ряді випадків неможливо, тому що для роботи апаратури часто потрібно забезпечити її електричний зв'язок із зовнішніми пристроями. Тому використовуються менш надійні засоби захисту, такі, як струмопровідні сітки, або плівкові покриття для вікон, щільникові металеві конструкції для повітрезабірників і вентиляційних отворів і контактні пружинні прокладки, розміщені по периметру дверей і люків.

Більш складною технічною проблемою рахується захист від проникнення ЕМІ в апаратуру через різноманітні кабельні входи. Радикальним рішенням даної проблеми міг би стати перехід від електричних мереж зв'язку до практично не схильних до впливу ЕМІ волоконно-оптичних. Проте заміна напівпровідникових приладів у всьому спектрі виконуваних ними функцій електронно-оптичними пристроями можлива тільки у віддаленому майбутньому. Тому в даний час в якості засобів захисту кабельних входів найбільші широко використовуються фільтри, у тому числі волоконні, а також іскрові розрядники, металлоокисні варистори та ін. [35].

Металлоокисні варистори є напівпровідниковими приладами, що різко підвищують свою провідність при високій напрузі. Проте, при застосуванні цих

приладів у якості засобів захисту від ЕМІ варто враховувати їх недостатньо високу швидкодію і погіршення характеристик при кількаразовому впливі навантажень. Ці недоліки відсутні у високошвидкісних зенеровських діодах, дія яких заснована на різкій лавиноподібній зміні опору від високого значення практично до нуля, при перевищенні прикладеної до них напруги граничного розміру. Крім того на відміну від варисторів характеристики зенеровських діодів після багатократних впливів високих напруг і переключень режимів не погіршуються.

Найбільш раціональним підходом до проектування засобів захисту від ЕМІ кабельних входів є створення таких роз'ємів у конструкції яких передбачені спеціальні заходи, що забезпечують формування елементів фільтрів і установку вмонтованих зенеровських діодів. Подібне рішення сприяє одержанню дуже малих значень ємності й індуктивності, що необхідно для забезпечення захисту від імпульсів, що мають незначну тривалість і, отже, потужну високочастотну складову.

4.3 Висновки до четвертого розділу

В цьому розділі розглянуто важливі питання охорони праці та безпеки в надзвичайних ситуаціях, зокрема описано закордонний досвід організації охорони праці в ІТ-компаніях та виконано оцінку дії на елементи комп'ютерної системи.

ВИСНОВКИ

У роботі було досягнуто мети та виконано поставлені завдання. Була розроблена архітектура моделі трансформера, що дозволяє враховувати взаємозв'язки між вимірюваннями, так і між локальними областями вздовж кожного вимірювання. Для досягнення цієї мети були зібрані дані для навчання та валідації моделі, підібрано універсальну архітектуру трансформера, проведено тестування розробленої моделі, а також розроблено інструмент автоматичного підбору гіперпараметрів моделі.

Для виконання поставлених завдань було проведено огляд існуючих рішень та з'ясовано, що моделі трансформерів показують найкращі результати порівняно з іншими моделями та підходами. Було виділено основні стратегії токенизації, які включають такі види: Time point token; Time variable token; Patching.

Аналізуючи підходи до вирішення задачі класифікації БЧР було виявлено недоліки цих методів. Було з'ясовано, що ці підходи не враховують зміни в локальних областях між вимірами. А також існуючі підходи включають складну передобробку і великовагову архітектуру, що сильно відбивається на швидкості навчання та інференсу моделей.

З метою усунення цих недоліків було розроблено архітектуру моделі трансформера для класифікації БЧР. Ключова ідея архітектури полягає у застосуванні механізму уваги як до послідовностей вимірів, так і послідовностей вздовж кожного виміру.

Для навчання та тестування розробленої моделі були висунуті вимоги до даних, які включає різні характеристики датасетів. Під ці вимоги ідеально підійшов набір даних UEA, на якому проводилася валідація розробленої моделі класифікації БЧР.

Для проведення коректного тестування архітектури було реалізовано інструмент автоматичного підбору гіперпараметрів. Даний інструмент дозволяє підібрати оптимальні параметри попередньої обробки та моделі для досягнення

найкращого результату. Цей інструмент був використаний для дотримання ідентичних умов проведення експериментів для всіх датасетів використовуваного набору.

Аналізуючи результати тестування, можна виявити, що розроблена модель точно випередила SOTA модель в 13 з 30 датасетів. А також отримала середню точність по всіх датасетах вище на 0,37%. Додатковою перевагою розробленої архітектури є те, що вона легковагова щодо SOTA моделі, що дозволяє істотно прискорити роботу моделі в режимі інференсу.

Для дослідження результатів тестування було вивчено основні особливості розробленої архітектури. Ключовими характеристиками моделі є розмір патчів і кількість енкодерів різних типів. Було виявлено, що із збільшенням розміру послідовностей необхідно зменшувати розмір патчів.

Перспективами продовження роботи є покращення результатів класифікації за допомогою використання патчів різної розмірності, використання потужніших пристроїв для збільшення кількості енкодерів, а також використання новітніх технік позиційного кодування часових рядів.

ПЕРЕЛІК ДЖЕРЕЛ

1. Christopher Chatfield. The Analysis of Time Series: Theory and Practice / Christopher Chatfield. New York: Springer-Science+Business Media, B.V., 2013. 263 с.
2. Abdullayeva F., Imamverdiyev Y. Development of oil production forecasting method based on deep learning //Statistics, Optimization & Information Computing. 2019. V. 7. №. 4. P. 826-839.
3. Liu X. et al. Deep learning in ECG diagnosis: A review //Knowledge-Based Systems. 2021. V. 227. P. 107187.
4. Assaad R. H., Fayek S. Predicting the price of crude oil and its fluctuations using computational econometrics: deep learning, LSTM, and convolutional neural networks //Econometric Research in Finance. 2021. V. 6. №. 2. P. 119-137.
5. Ozbayoglu A. M., Gudelek M. U., Sezer O. B. Deep learning for financial applications: A survey //Applied Soft Computing. 2020. V. 93. P. 106384.
6. Jansen S. Machine Learning for Algorithmic Trading: 2020. [Електронний ресурс] – Режим доступу: <https://www.amazon.com/Machine-Learning-Algorithmic-Tradingalternative/dp/1839217715> (дата звернення: 18.11.2025).
7. Leveraging XGBoost for Time-Series Forecasting. [Електронний ресурс] – Режим доступу <https://www.kdnuggets.com/2023/08/leveraging-xgboost-timeseries-forecasting.html> (дата звернення: 19.11.2025).
8. Берзлев О. Ю. Сучасний стан інформаційних систем прогнозування часових рядів / О. Ю. Берзлев // ДВНЗ «Ужгородський національний університет». 2013. №13. С. 78–82.
9. Кащєєва В. Ю. Інформаційна технологія аналізу багатокomпонентних процесів за часовими рядами на основі інтервальних прогнозних моделей : автореф. дис. ... канд. техн. наук : 05.13.06. Харків, 2014.
10. Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. Journal of Neural Engineering, 16(5), 051001.

11. Lai R. K. et al. Evolving and clustering fuzzy decision tree for financial time series data forecasting //Expert Systems with Applications. 2009. V. 36. №. 2. P. 3761-3773.
12. Chen Y. et al. Dtw-d: time series semi-supervised learning from a single example //Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. 2013. P. 383-391.
13. Siami-Namini S., Tavakoli N., Namin A. S. The performance of LSTM and BiLSTM in forecasting time series //2019 IEEE International conference on big data (Big Data). IEEE, 2019. P. 3285-3292.
14. Wang Z., Yan W., Oates T. Time series classification from scratch with deep neural networks: A strong baseline //2017 International joint conference on neural networks (IJCNN). IEEE, 2017. P. 1578-1585.
15. Yue Z. et al. Ts2vec: Towards universal representation of time series //Proceedings of the AAAI Conference on Artificial Intelligence. 2022. V. 36. №. 8. P. 8980-8987.
16. Vaswani, A., et al. Attention Is All You Need. Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, 4-9 December 2017, pp. 6000-6010
17. Yun C. et al. Are transformers universal approximators of sequence-to-sequence functions? //arXiv preprint arXiv:1912.10077. 2019.
18. Zhou H. et al. Informer: Beyond efficient transformer for long sequence timeseries forecasting //Proceedings of the AAAI conference on artificial intelligence. 2021. V. 35. №. 12. P. 11106-11115
19. Wu H. et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting //Advances in Neural Information Processing Systems. 2021. V. 34. P. 22419-22430.
20. Zhou T. et al. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting //International Conference on Machine Learning. – PMLR, 2022. P. 27268-27286

21. Zeng A. et al. Are transformers effective for time series forecasting? //Proceedings of the AAAI conference on artificial intelligence. 2023. V. 37. №. 9. P. 11121-11128.
22. Nie Y. et al. A time series is worth 64 words: Long-term forecasting with transformers //arXiv preprint arXiv:2211.14730. 2022.
23. Yu C. et al. DSformer: a double sampling transformer for multivariate time series long-term prediction //Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. 2023. P. 3062-3072
24. Wen Q. et al. Transformers in time series: A survey //arXiv preprint arXiv:2202.07125. 2022, V.5, P.19-27.
25. Zhou Z. et al. A K-variate Time Series Is Worth K Words: Evolution of the Vanilla Transformer Architecture for Long-term Multivariate Time Series Forecasting //arXiv preprint arXiv:2212.02789. 2022.
26. Bagnall A. et al. The UEA multivariate time series classification archive, //arXiv preprint arXiv:1811.00075. 2018
27. Zerveas G. et al. A transformer-based framework for multivariate time series representation learning //Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining. 2021. P. 2114-2124.
28. Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding //arXiv preprint arXiv:1810.04805. 2018.
29. Liu M. et al. Gated transformer networks for multivariate time series classification //arXiv preprint arXiv:2103.14438. 2021.
30. Baydogan M. G., Runger G. Learning a symbolic representation for multivariate time series classification //Data Mining and Knowledge Discovery. 2015. V. 29. P. 400-422.
31. Chowdhury R. R. et al. Tarnet: Task-aware reconstruction for time-series transformer //Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022. P. 212-220.
32. Yang C. et al. Dyformer: A dynamic transformer-based architecture for multivariate time series classification //Information Sciences. 2024. V. 656.P. 119881.

33. Yang C. H. H., Tsai Y. Y., Chen P. Y. Voice2series: Reprogramming acoustic models for time series classification //International conference on machine learning. PMLR, 2021. P. 11808-11819.
34. Rußwurm M., Körner M. Self-attention for raw optical satellite time series classification //ISPRS journal of photogrammetry and remote sensing. 2020. V. 169. P. 421-435
35. Webster J. J., Kit C. Tokenization as the initial phase in NLP //COLING 1992 volume 4: The 14th international conference on computational linguistics. 1992.
36. Su L. et al. Beyond Two-Tower Matching: Learning Sparse Retrievable CrossInteractions for Recommendation //Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2023. p. 548-557.
37. Лісовий В.Г. Архітектура трансформера для класифікації багатовимірних часових рядів // XIV Міжнародна науково-практична конференція молодих учених та студентів «Актуальні задачі сучасних технологій», Тернопіль, 11-12. Грудня 2025 р. с. 289 – 291.
38. Лісовий В.Г., Семенишин Г.М. Алгоритм роботи програмного інструменту автоматичного підбору гіперпараметрів трансформера // Інформаційні моделі, системи та технології: Праці XIII наук.-техн. конф. Тернопіль, 17-18 грудня 2025 р. с. 56 - 57.
39. Python Documentation. [Електронний ресурс] – Режим доступу: <https://www.python.org/> (дата звертання: 16.11.2025).
40. Gradio Documentation [Електронний ресурс] – Режим доступу: <https://www.gradio.app/> (дата звертання: 17.11.2025).
41. Zhou T. et al. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting //International Conference on Machine Learning. PMLR, 2022. P. 27268-27286.
42. Yatsyshyn V., Pastukh O., Palamar A., Zharovsky R. Technology of relational database management systems performance evaluation during computer

systems design. Scientific Journal of TNTU, Ternopil, Ukraine, 2023. Vol. 109, No 1. P. 54–65.

43. Skorenkyu Y., Kozak R., Zagorodna N., Kramar O., Baran, I. Use of augmented reality-enabled prototyping of cyber-physical systems for improving cybersecurity education. Journal of Physics: Conference Series. 2021. Vol. 1840, No. 1. DOI: <http://dx.doi.org/10.1088/1742-6596/1840/1/012026>.

44. Микитишин А. Г., Митник М. М., Стухляк П. Д., Пасічник В. В. Комп'ютерні мережі. Книга 2. [навчальний посібник]. Львів : "Магнолія 2006", 2014. 312 с.

45. Микитишин А. Г., Митник М. М., Стухляк П. Д. Телекомунікаційні системи та мережі. Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2017. 384 с.

46. Сьогодні UA [Електронний ресурс] – Режим доступу: <https://www.segodnya.ua/lifestyle/fun/pochti-kak-u-googlechemudivlyayut-ofisy-ukrainskih-it-kompaniy--764025.html> (дата звертання 10.12.2025).

47. Як працюють в Google? Умови, в яких хочеться трудитися. [Електронний ресурс] – Режим доступу: <http://www.clevers.com.ua/articles-cleveradvertising-agency/success-stories/245-google2> (дата звертання 10.12.2025).

48. Офіс мрії: Робота в компанії Google. [Електронний ресурс] – Режим доступу: <http://bigpicture.ua/?p=187580> (дата звертання 11.12.2025).

49. Офіс Facebook: Репортаж із RMA SiliconTrip. [Електронний ресурс] – Режим доступу: <https://habrahabr.ru/company/rma/blog/103800/> (дата звертання 11.12.2025).

50. Класифікатор професій ДК 003:2010. [Електронний ресурс] – Режим доступу: <https://zakon.rada.gov.ua/rada/show/va327609-10> (дата звертання 11.12.2025).

51. Закон України «Про охорону праці». [Електронний ресурс] – Режим доступу: <https://zakon.rada.gov.ua/laws/show/2694-12> (дата звертання 12.12.2025).

52. ДНАОП 0.00-8.20-99. Порядок проведення експертизи електроустановок споживачів/ [Електронний ресурс] – Режим доступу:

https://dnaop.com/html/43255/doc-%D0%94%D0%9D%D0%90%D0%9E%D0%9F_0.00-8.20-99 - (дата звертання 12.12.2025).

53. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання / укл.: Стручок В. С. Тернопіль: ФОП Паляниця В. А., 2022. 156 с.

ДОДАТКИ

ДОДАТОК А

Тези конференції

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Тернопільський національний технічний університет імені Івана Пулюя
Маріборський університет (Словенія)
Технічний університет у Кошице (Словаччина)
Вільнюський технічний університет ім. Гедимінаса (Литва)
Краківський економічний університет (Польща)
Вроцлавський економічний університет (Польща)
Університет «Опольська Політехніка» (Польща)
Національний університет «Полтавська політехніка імені Юрія Кондратюка»
Вінницький національний аграрний університет
Львівський національний університет ім. І. Франка
Головне управління Пенсійного фонду в Тернопільській області
Наукове товариство ім. Шевченка
Тернопільський обласний комунальний інститут післядипломної педагогічної освіти
Сумський державний педагогічний університет
Запорізький національний університет

АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ

Збірник

тез доповідей

**XIV Міжнародної науково-технічної
конференції молодих учених та студентів**

11-12 грудня 2025 року



УКРАЇНА
ТЕРНОПІЛЬ – 2025

27.	О. Карнаухов, Т. Лобур, С. Мариенко АВТОМАТИЗОВАНА СИСТЕМА КОНТРОЛЮ І ОБЛІКУ ЕНЕРГОРЕСУРСІВ УНІВЕРСИТЕТУ	270
28.	В. Ю. Кашин, Т.А. Лечаченко АНАЛІЗ ВРАЗЛИВОСТЕЙ HTTP REQUEST SMUGGLING ЗАСОБАМИ ДИФЕРЕНЦІАЛЬНОГО ФАЗЗИНГУ HTTP-ЗАПИТІВ	272
29.	С.Б. Кіт, В.Р. Перун, С.І. Перун, Г.В. Шимчук ІНТЕЛЕКТУАЛЬНА СИСТЕМА КОНТРОЛЮ ПАРАМЕТРІВ МІКРОКЛІМАТУ В ЖИТЛОВОМУ ПРИМІЩЕННІ З ВИКОРИСТАННЯМ БСМ	273
30.	С.Б. Кіт, В.Р. Перун, С.І. Перун, Г.В. Шимчук ХМАРНО-ОРІЄНТОВАНА ПЛАТФОРМА СПОСТЕРЕЖЕННЯ ЗА МІКРОКЛІМАТОМ КВАРТИРИ НА ОСНОВІ ІОТ	275
31.	А.М. Ковтко, І.Р. Козбур, В.Б. Савків, Г.В. Козбур АРХІТЕКТУРА АВТОМАТИЗОВАНОЇ СИСТЕМИ ДЛЯ ГЕНЕРАЦІЇ МОДУЛЬНИХ ТЕСТІВ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ ТА МУТАЦІЙНОГО ТЕСТУВАННЯ	278
32.	К. А. Кокурін РОЗРОБКА СИСТЕМИ МОНІТОРИНГУ ЯКОСТІ ПОВІТРЯ З ВИКОРИСТАННЯМ СЕНСОРА SDS011 ТА TELEGRAM-БОТА ДЛЯ СПОВІЩЕНЬ	280
33.	М.А. Конотопський, Ю.З. Лещини МЕТОДИ ТА ПРОГРАМНО-АПАРАТНІ ЗАСОБИ КОМП'ЮТЕРИЗОВАНОГО КЕРУВАННЯ ПОТУЖНІСТЮ ТЕПЛОПОСТАЧАЛЬНОГО ПУНКТУ	282
34.	В.О. Кравчик МЕТОДИ ТА ЗАСОБИ ВИЯВЛЕННЯ ДОРОЖНЬО-ТРАНСПОРТНИХ ПРИГОД КОМП'ЮТЕРИЗОВАНИМИ СИСТЕМАМИ ВІДЕОНАГЛЯДУ	283
35.	В.В. Левіцький, А. Г. Микитшин, А.Ю. Гураль, А.В. Михайлюк АВТОМАТИЗОВАНА СИСТЕМА КЕРУВАННЯ ТЕХНОЛОГІЧНИМ ПРОЦЕСОМ ВИГОТОВЛЕННЯ КИСЛОМОЛОЧНИХ СИРІВ	284
36.	Р.В. Лесик ДОСЛІДЖЕННЯ АДАПТИВНИХ АЛГОРИТМІВ РЕГУЛЮВАННЯ КОГНІТИВНОГО НАВАНТАЖЕННЯ У ВЕБТРЕНАЖЕРІ ПАМ'ЯТІ	286
37.	Т.А. Ліпак ПРИНЦИПИ ПРОЄКТУВАННЯ МУЛЬТИМОДАЛЬНИХ ІНТЕРФЕЙСІВ З МОБІЛЬНОЮ ДОПОВНЕНОЮ РЕАЛЬНОСТЮ В ІОТ	288
38.	В.Г. Лісовий АРХІТЕКТУРА ТРАНСФОРМЕРА ДЛЯ КЛАСИФІКАЦІЇ БАГАТОВИМІРНИХ ЧАСОВИХ РЯДІВ	289
39.	М. В. Лісовий, Г. І. Ліпак ПРОЄКТУВАННЯ ЕЛАСТИЧНИХ ХМАРНИХ АРХІТЕКТУР ДЛЯ СТІЙКОСТІ ЦИФРОВИХ СЕРВІСІВ ГРОМАДСЬКОЇ ВЗАЄМОДІЇ	292
40.	А.М. Луцків, Д.М. Гаврада АРХІТЕКТУРА АПАРАТНО-ПРОГРАМНОГО КОМПЛЕКСУ МУЛЬТИМОДАЛЬНОЇ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ ОСОБИ	293
41.	А. Луцків, С. Андруньків РОЗРОБКА СИСТЕМИ МОНІТОРИНГУ СЕМАНТИЧНОЇ ЯКОСТІ ТА ГАЛЮЦИНАЦІЙ ДЛЯ МОДЕЛЕЙ LLM В MLOPS	295
42.	А.М. Луцків, В.В. Комарницький DEVOPS-ПІДХІД ДО АВТОМАТИЗАЦІЇ CI/CD У РОЗПОДІЛЕНИХ ІОТ-СИСТЕМАХ	296

УДК 004.8

В.Г. Лісовий – студент гр. САм-61

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

АРХІТЕКТУРА ТРАНСФОРМЕРА ДЛЯ КЛАСИФІКАЦІЇ БАГАТОВИМІРНИХ ЧАСОВИХ РЯДІВ

V. Lisovyi st.group SAm-61

TRANSFORMER ARCHITECTURE FOR MULTIDIMENSIONAL TIME SERIES CLASSIFICATION

Для вирішення задачі класифікації багатовимірних часових рядів розроблено архітектуру моделі трансформера, що дозволяє враховувати як взаємозв'язки між локальними областями серед вимірів, так і взаємозв'язки між підпоследовностями вздовж кожного виміру. Як основа взята архітектура двобаштової моделі [1], яка може обробляти різну природу даних. Розроблена архітектура показано на рис. 1.

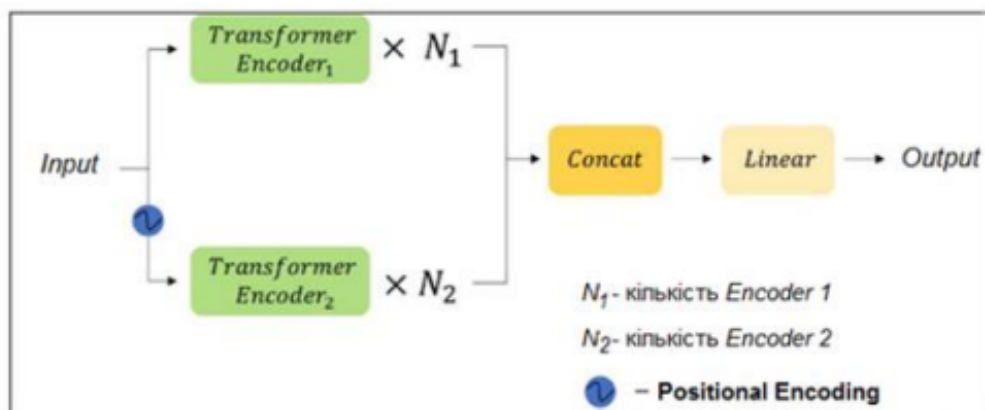


Рисунок 1 – Архітектура моделі, що розробляється

Вхідні дані надходять у два різних кодувальники, котрі називаються Encoder 1 і Encoder 2 відповідно, додаючи у другому випадку позиційний вектор. Далі дані проходять через N шарів кодувальників, що об'єднуються, проходять через лінійний шар і формується цільовий результат.

Як вхідні дані використовуються багатовимірні часові ряди, розділені на підпоследовності однакової розмірності (патчі). При проходженні даних через розроблену модель формується набір ймовірностей, що дозволяє визначити клас найбільшої ймовірності.

Як кодувальники використовуються оригінальні шари трансформер моделі, представлені в роботі [2]. Дані шари перетворюють вхідну послідовність у внутрішнє представлення, що містить інформацію про семантику та структуру вихідних даних. Він складається з декількох блоків, кожен з яких містить безліч шарів самоуваги (self-attention) і шарів прямого (feed-forward). Перші шари допомагають моделі фокусуватися на різних частинах вхідних даних, а другі шари забезпечують нелінійні перетворення, покращуючи здатність моделі до вивчення складних залежностей даних. Кількість кожного кодувальника може змінюватись і є гіперпараметром моделі. Завданням Encoder 1 є пошук закономірностей між патчами серед вимірів. А завданням Encoder 2 є пошук закономірностей між патчами вздовж кожного часового ряду. Пошук взаємозв'язків кодувальниками показано на рис. 2.

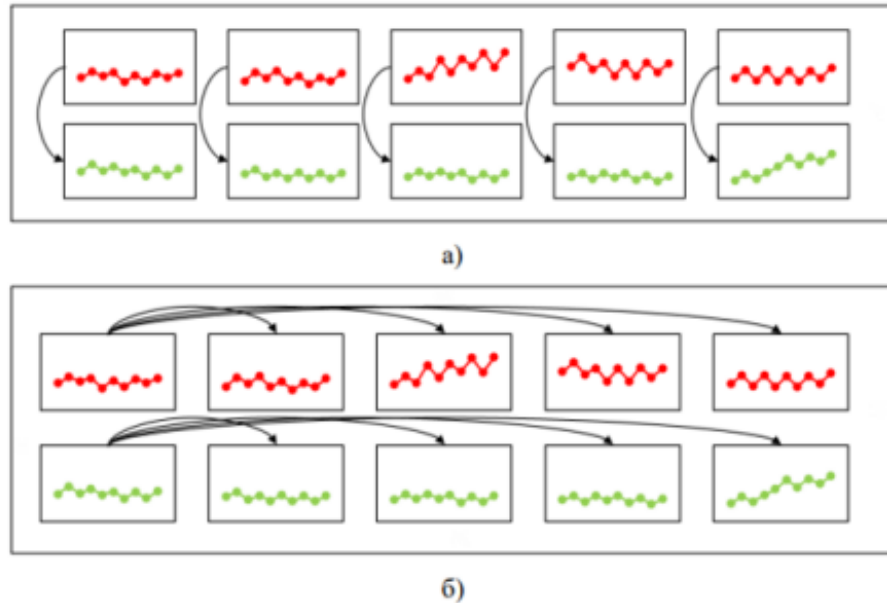


Рисунок 2 – Пошук закономірностей: а) Encoder 1; б) Encoder2

Основна ідея даної архітектури полягає у пошуку закономірностей між вимірами, що дозволяє досягти Encoder 1. Оскільки кодувальники не враховують упорядкованість послідовностей, прийнято додавати позиційний вектор до даних. Але оскільки серед вимірів немає порядку, цей вектор додається тільки до даних, які проходять через Encoder 2.

Як стратегія позиціонування використовується оригінальний метод з роботи [2], який полягає у використанні тригонометричних функцій для отримання позиційного вектора. Позиційне кодування дозволяє враховувати абсолютні та відносні позиції патчів вздовж кожного виміру.

Для оцінки розробленої архітектури завдання класифікації багатовимірних часових рядів було встановлено критерії відбору даних. Вибрані дані повинні представляти різні сфери та мати різноманітну природу. Крім того, вони повинні мати різні розміри вхідних і вихідних даних, а також бути актуальними, зведеними до однієї довжини і попередньо очищеними від шумів і аномалій.

Література

1. Su L. et al. Beyond Two-Tower Matching: Learning Sparse Retrievable CrossInteractions for Recommendation //Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2023. pp. 548-557.
2. Vaswani, A., et al. Attention Is All You Need. Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, 4-9 December 2017, pp. 6000-6010

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

МАТЕРІАЛИ

ХІІ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

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



17-18 грудня 2025 року

**ТЕРНОПІЛЬ
2025**

КІБЕРЗАГРОЗ У МЕДИЧНИХ СИСТЕМАХ

D. Kosaryk, R. Kozak

APPLICATION OF THE MONTE CARLO METHOD FOR MODELING
CYBERTHREATS IN MEDICAL SYSTEMS

Т. Крамар, П. Скалецький

ПОРІВНЯННЯ ІНФОРМАЦІЙНО-ТЕХНОЛОГІЧНИХ ПЛАТФОРМ ДЛЯ
ЗБЕРЕЖЕННЯ ТА ПОДАВАННЯ КУЛЬТУРНОЇ СПАДЩИНИ

T. Kramar, P. Skaletskyi

COMPARISON OF INFORMATION-TECHNOLOGICAL PLATFORMS FOR THE
PRESERVATION AND PRESENTATION OF CULTURAL HERITAGE

52

О. Легкобйт, М. Стадник

ГЕНЕРАТИВНІ МОВНІ МОДЕЛІ В АНАЛІЗІ ШКІДЛИВОГО КОДУ

O. Lehkobyt, M. Stadnyk

GENERATIVE LANGUAGE MODELS IN MALWARE ANALYSIS

53

А. Луцків, С. Андруньків

ІНТЕРПРЕТАЦІЯ 'ЧОРНИХ СКРИНЬОК'. ВИКОРИСТАННЯ МЕТОДІВ SHAP
ТА LIME ДЛЯ ВІЗУАЛІЗАЦІЇ ПРОЦЕСУ ПРИЙНЯТТЯ РІШЕНЬ У ГЛИБОКИХ
НЕЙРОННИХ МЕРЕЖАХ

A. Lutskev, S. Andrunkev

INTERPRETATION OF 'BLACK BOXES'. USE OF SHAP AND LIME METHODS
TO VISUALIZE THE DECISION-MAKING PROCESS IN DEEP NEURAL
NETWORKS

54

Г. Липак, І. Генсьора

УПРАВЛІННЯ ЦИФРОВИМИ КОЛЕКЦІЯМИ МУЗЕЇВ: ЗАСОБИ І СТАНДАРТИ

H. Lyapak, I. Hensora

MANAGEMENT OF DIGITAL MUSEUM COLLECTIONS: TOOLS AND
STANDARDS

55

В. Лісовий, В. Яцишин, Г. Семеншин

АЛГОРИТМ РОБОТИ ПРОГРАМНОГО ІНСТРУМЕНТУ АВТОМАТИЧНОГО
ПІДБОРУ ГІПЕРПАРАМЕТРІВ ТРАНСФОРМЕРА

V. Lisovyi, V. Yatsyshyn, H. Semenishyn

ALGORITHM OF THE SOFTWARE TOOL FOR AUTOMATIC SELECTION OF
TRANSFORMER HYPERPARAMETERS

56

С. Літвінчук, Н. Загородна

ДОСЛІДЖЕННЯ БЕЗПЕРЕРВНОЇ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ НА
ОСНОВІ ДИНАМІКИ НАТИСКАННЯ КЛАВІШ

S. Litvinchuk, N. Zagorodna

RESEARCH OF CONTINUOUS USER AUTHENTICATION BASED ON
KEYSTROKE DYNAMICS

58

Д. Мазурчак, В. Тимошук

ДОСЛІДЖЕННЯ МЕХАНІЗМУ ІЗОЛЯЦІЇ ПАМ'ЯТІ ДЛЯ ПІДВИЩЕННЯ
БЕЗПЕКИ ВІРТУАЛІЗОВАНИХ МЕРЕЖЕВИХ СЕРЕДОВИЩАХ

D. Mazurchak, V. Tymoshchuk

INVESTIGATION OF MEMORY ISOLATION TECHNIQUES FOR IMPROVING
THE SECURITY OF VIRTUALIZED NETWORK SYSTEMS

59

К. Марущак, Г. Козбур

ТИПОЛОГІЯ ІНСТРУМЕНТІВ ВІЗУАЛІЗАЦІЇ ДАНИХ В СУЧАСНІЙ
АНАЛІТИЦІ

K. Marushchak, Halyna Kozbur

TYPOLOGY OF DATA VISUALIZATION TOOLS IN MODERN ANALYTICS

60

Д. Марчук, В. Тимошук

АРХІТЕКТУРА БЕЗПЕКИ ЯДРА 5G НА ОСНОВІ ВІРТУАЛІЗОВАНИХ
МЕРЕЖЕВИХ ФУНКЦІЙ ТА УЗГОДЖЕНИХ ПОЛІТИК

61

УДК 004.8

В. Лісовий; В. Яцишин к. т. н., доц.; Г. Семеншин

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

АЛГОРИТМ РОБОТИ ПРОГРАМНОГО ІНСТРУМЕНТУ АВТОМАТИЧНОГО ПІДБОРУ ГІПЕРПАРАМЕТРІВ ТРАНСФОРМЕРА

UDC 004.8

V. Lisovyi; V. Yatsyshyn PhD., Assoc. Prof.; H. Semenyshyn

ALGORITHM OF THE SOFTWARE TOOL FOR AUTOMATIC SELECTION OF TRANSFORMER HYPERPARAMETERS

Для автоматичного підбору найкращих параметрів передоброби та гіперпараметрів розробленої моделі класифікації багатовимірних часових рядів (БЧР) розроблено модуль автоматичного перебору різних параметрів. Вхідні дані навчання та тестування приймаються у модуль у форматі датасетів UEA [1].

Цей модуль реалізований мовою програмування Python [2] та виконує послідовний перебір параметрів для досягнення найкращих результатів. Основна ідея полягає у переборі кожного гіперпараметра та збереження значення, за допомогою якого було отримано найбільшу метрику на контрольній вибірці. Як метрика оцінки використана частка правильних відповідей серед усіх прикладів.

Початкові значення гіперпараметрів та його діапазони перебору представляються у файлі конфігурації. Цей файл приймається у форматі JSON. Також у файлі конфігурації представлені обмеження гіперпараметрів для ефективного підбору при вирішенні завдань з більшими обсягами даних.

Далі представлений алгоритм виконання автоматичного перебору параметрів передоброби та гіперпараметрів моделі.

1. Ініціалізація початкових параметрів. На даному етапі відбувається встановлення початкових значень параметрів та гіперпараметрів, які будуть використовуватись у подальшому процесі перебору. Ініціалізація початкових параметрів потрібна для того, щоб задати відповідні точки для подальшого перебору.

2. Перебір параметрів нормалізації. На цьому етапі перебираються різні методи нормалізації даних. Використовується нормалізація за мінімумом та максимумом. Проводиться перебір таких варіантів як нормування кожного виміру окремо, нормування щодо глобального мінімуму та максимуму, не використання нормалізації. Нормалізація даних допомагає покращити збіжність алгоритмів навчання та збільшити точність моделі.

3. Використання перетворення Фур'є. На цьому етапі розглядається застосування перетворення Фур'є до даних, що дозволяє перейти від часової до частотної області. Перетворення Фур'є може виявити важливі частотні компоненти даних, які можуть бути корисні для поліпшення якості моделі.

4. Перебір варіантів об'єднання та вибору вимірювань. Перебираються різні стратегії об'єднання та вибору вимірів, такі як об'єднання всіх вимірів в один масив, вибірка найбільш значимих вимірів та інші підходи. Правильний вибір та поєднання вимірів може значно вплинути на продуктивність моделі. Це допомагає зменшити розмірність даних та покращити ефективність навчання.

5. Перебір розміру патчів. На цьому етапі перебираються різні розміри патчів, які будуть використовуватися при розбитті даних на дрібніші частини. Розмір патчів може впливати на здатність моделі захоплювати локальні патерни даних. Правильний вибір розміру патчів може покращити навчання моделі та її здатність до узагальнення. Розміри для перебору залежать від довжини вхідної послідовності та є дільниками довжини.

6. Перебір кроку навчання. Перебираються різні значення кроку навчання, який визначає, наскільки модель оновлює свої ваги при кожному кроці навчання. Крок навчання є одним із ключових гіперпараметрів, який впливає на збіжність та стабільність процесу

навчання. Правильний крок навчання допомагає уникнути як застрягання в локальних мінімумах, так і перескакування через оптимальні значення. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 0.00001, 0.0001, 0.001, 0.01.

7. Перебір кількості нейронів у повнозв'язкових шарах енкодерів. Правильний вибір цього параметра допомагає досягти балансу між складністю моделі та її здатністю до узагальнення. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 516, 1024, 2048.

8. Перебір частки занулення нейронів (Dropout). Перебираються різні частки занулення нейронів в енкодер шарах моделі. Даний гіперпараметр дозволяє зменшити перенавчання та підвищити стійкість моделі. Можливі значення для перебору можуть встановлюватись користувачем. У роботі були використані значення 0.1, 0.25, 0.5, 0.75.

9. Перебір кількості енкодерів. Перебираються різні кількості енкодерів, що використовуються в моделі. Кількість енкодерів впливає на глибину моделі та її здатність захоплювати складні патерни у даних. Перебір дозволяє знайти оптимальну кількість енкодерів для вирішення задачі. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 0, 1, 2, 3.

10. Використання шару ембедингів. Розглядається включення чи виключення шару ембедингів до архітектури моделі. Ембединги можуть поліпшити представлення даних моделі і це може підвищити точність моделі.

11. Перебір кількості голів уваги. Перебираються різні кількості голів уваги моделі. Кількість голів уваги визначає, скільки різних аспектів інформації модель може одночасно враховувати. Це може значно покращити здатність моделі до навчання та її продуктивність на складних завданнях. Можливі значення для перебору можуть встановлюватись користувачем. У роботі було використано значення 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

Перебір вищевикладених гіперпараметрів у представлений послідовності дозволяє досягти найкращих результатів для вирішення задачі класифікації БЧР за допомогою розробленої моделі трансформера.

Література

1. Bagnall A. et al. The UEA multivariate time series classification archive, 2018 //arXiv preprint arXiv:1811.00075.
2. Python Documentation // Python URL: <https://www.python.org/> (дата звертання: 24.11.2025).