

9. Мигаль В.Д., Аргун, Щ. В., Гнатов, А. В., Гнатова, Г. А., & Сохін, П. А. (2023) Підвищення якості тягових електродвигунів для електротранспорту. Ав-томобіль і електроніка. Сучасні технології: електронне наукове спеціалізоване видання. (23), С. 6–14. DOI:10.30977/VEIT.2023.23.0.1

10. Hnatov, A., & Arhun, S. (2022). Electric vehicles and energy-saving technologies – master’s degree program under the Erasmus project Cybphys. Automobile Transport, (51), 85–95.

МОДУЛЬНА АРХІТЕКТУРА ОБРОБКИ ТА ЗБЕРІГАННЯ ДАНИХ ТРАНСПОРТНИХ ЗАСОБІВ

Грицук Валерій Юрійович, аспірант, Харківський національний автомобільно-дорожній університет, e-mail: valeri.gritsuk@gmail.com,
ORCID: <https://orcid.org/0000-0002-3780-7815>

Грицук Юрій Валерійович, к.т.н., доцент, Донбаська національна академія будівництва і архітектури, e-mail: yuri.gritsuk@gmail.com,
ORCID: <https://orcid.org/0000-0003-3389-1172>

Сучасні транспортні засоби генерують величезну кількість даних [1-3] завдяки вбудованим сенсорам, що забезпечують моніторинг показників роботи двигуна, температурних режимів, кодів несправностей та інших критичних параметрів. Для ефективної діагностики та прогнозування можливих відмов було розроблено інтегровану систему, яка забезпечує збирання, попередню обробку, зберігання та аналіз даних у режимі реального часу. Основні етапи включають зчитування даних за допомогою OBD-II сканера, їх перевірку на цілісність і достовірність, нормалізацію, подальшу передачу на внутрішній сервер, організацію у базі даних (як поточних, так і історичних) з наступною синхронізацією, а також аналіз за допомогою нейронної мережі, яка виявляє закономірності та аномалії. Використання сучасних технологій (Apache Spark, Kafka, контейнеризації, мікросервісів) та зовнішніх бібліотек машинного навчання (TensorFlow, PyTorch, scikit-learn та ін.) забезпечує відмовостійкість, масштабованість та високу пропускну здатність системи.

На діаграмі активності системи моніторингу параметрів транспортного засобу [2] (рис. 1) можна побачити взаємодію компонентів між собою. Сканер OBD-II в режимі реального часу зчитує дані з датчиків, зокрема показники роботи двигуна, коди несправностей, температури компонентів та інші. Внутрішній препроцесор перевіряє дані на цілісність та достовірність, попередньо відфільтровуючи очевидні помилки або відсутні значення.

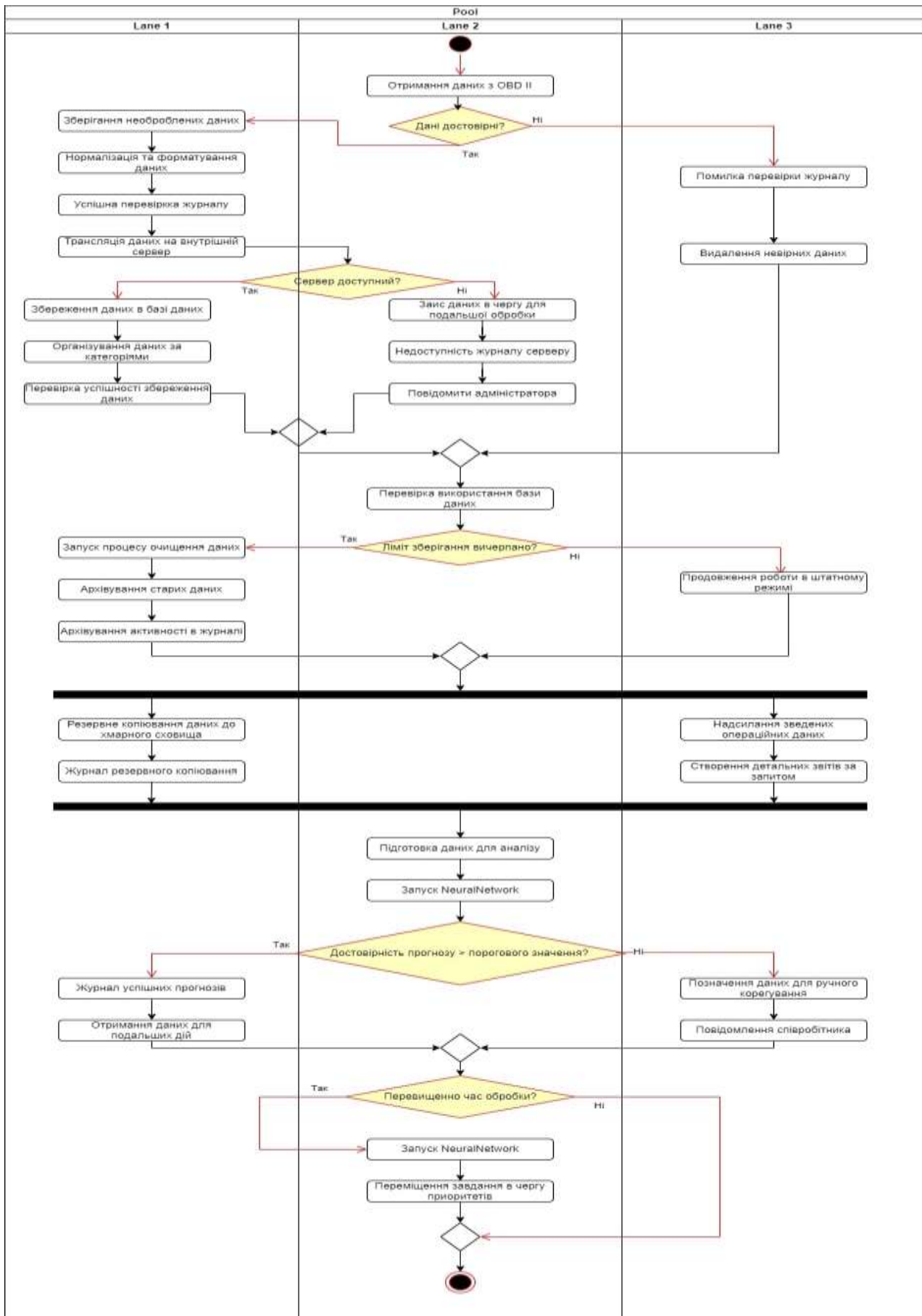


Рисунок 1 – Діаграма активності системи

Якщо дані недостовірні або не повні, в журнал реєструється помилка, і дані або повторно запитуються у автомобіля, або позначаються для ручного перегляду, або позначаються на видалення. При умові «Так» далі виконується перевірка цілісності та наявності всіх очікуваних даних. Також проходить перевірка показників датчиків на предмет відповідності очікуваним робочим діапазнам. Нормалізатор даних конвертує сирі дані датчика в стандартизований формат для подальшої обробки. Оброблені дані надсилаються на внутрішній сервер для зберігання та аналізу. Разом з цим, починається паралельна обробка даних.

Під час зберігання дані також передаються до сховища даних для довготривалого зберігання. Внутрішній сервер здійснює зберігання оброблених і перевічених даних на сервері бази даних для негайного доступу до них. При збереженні даних, сервер організовує дані за категоріями (наприклад, «поточні» та «історичні»). Паралельно система резервного копіювання надсилає копії важливих даних до хмарного сховища для аварійного відновлення. Сховище та база даних перевіряють інформацію та видаляють надлишкові або застарілі записи.

Дані, історичні та поточні, проходять до нейронної мережі в реальному часі з сервера сховища даних. Модель машинного навчання аналізує структуру цих даних для виявлення закономірностей та виявлені аномалій, а саме появу нових дефектів або нестандартну поведінку датчиків.

Для сучасних платформ незамінними є зовнішні бібліотеки, які дозволяють реалізувати вже готові функції, такі як машинне навчання, робота з базами даних та обробка даних. Однак їх інтеграція передбачає необхідність чіткого архітектурного планування для забезпечення сумісності системи, відмовостійкості, масштабованості та дотримання вимог безпеки. Відповідно, для коректної роботи машинного навчання потрібно підключити відповідні бібліотеки (рис. 2), які використовують алгоритми та моделі машинного навчання для прогностичного аналізу (наприклад, TensorFlow, PyTorch, scikit-learn [4, 5], ONNX Runtime та інші). Це дозволить тренувати індивідуальні моделі, використовуючи різні набори даних та підтримувати зворотнє навчання для покращення для конкретної області. В систему додаються процеси, такі як мікросервіси, що працюють у контейнерах (наприклад, Docker). Для цього використовується спеціальний сервер виведення (наприклад, NVIDIA Triton, TensorFlow Serving). Це дозволяє зменшити кількість дублікатів і забезпечити узгоджену взаємодію між модулями, а також дозволяє використовувати інтерфейси API для виведення в реальному часі або пакетного формування висновків. Водночас, з метою проведення аналітики в реальному часі в умовах низьких затримок, легкі моделі можна розгнати на периферійних пристроях (наприклад, OBD-II обладнання з вбудованим TensorFlow Lite [6, 7]).

Бібліотеки машинного навчання легко поєднується з бібліотекою обробки даних для попередньої обробки даних, що забезпечує нормалізацію та готовність цих даних до аналізу. Крім того, бібліотека взаємодіє з модулем очищення даних для забезпечення якості вхідних даних, що прямо впливає на точність моделі.

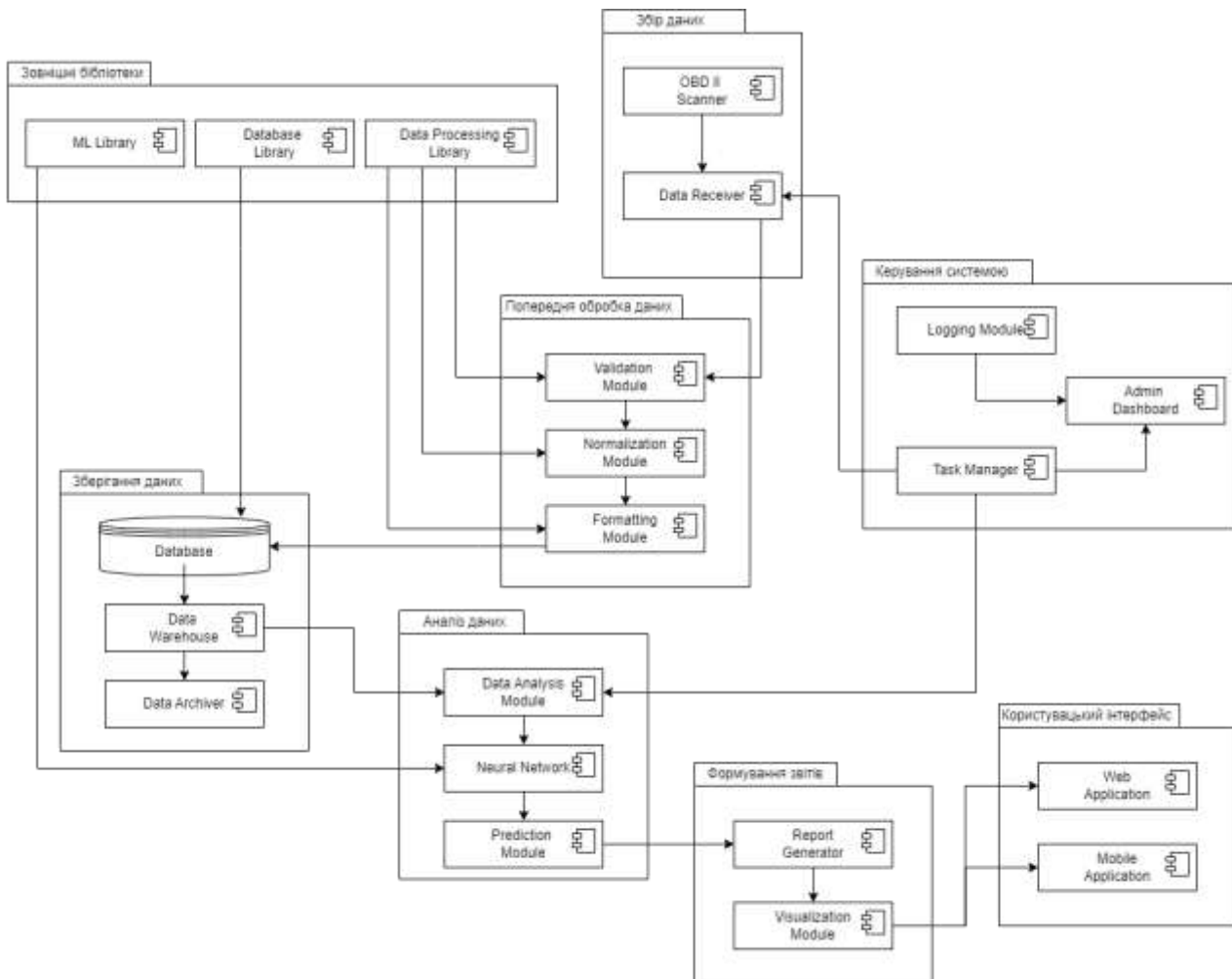


Рисунок 2 – Діаграма компонентів системи

При пакетній обробці відсіювання недійсних записів, агрегування даних з датчиків та нормалізацію одиниць виміру. Також відбувається видалення записів з відсутніми критично важливими полями (наприклад, ідентифікатор транспортного засобу або мітка часу) та перевірка діапазонів. При нормалізації стандартизуються одиниці вимірювання та відбувається перетворення різних протоколів OBD-II (наприклад, ISO 15765-4, SAE J1850) в уніфікований формат JSON.

Також разом з цим відбуваються обчислення окремих характеристик, такі як паливна ефективність та інші, групуються за ідентифікатором транспортного засобу та узагальнюються метрики. Далі попередньо оброблені дані

надсилаються до валідатора даних для додаткових перевірок. І в кінці очищені дані спрямовуються до бази даних для довготривалого зберігання. Apache Spark та Kafka ефективно використовують розбиття даних на блоки для паралельного виконання.

Розроблена система демонструє успішну інтеграцію різних компонентів, що забезпечують збирання даних від датчиків, їх попередню перевірку, нормалізацію та зберігання як у режимі реального часу, так і для довгострокового аналізу. Інтеграція нейронних мереж та використання сучасних бібліотек машинного навчання дозволяє ефективно аналізувати поточні та історичні дані, виявляючи закономірності та аномальні поведінки, що є ключовим для своєчасного прогнозування дефектів. Завдяки використанню мікросервісів, контейнеризації та розподіленої обробки даних, система забезпечує високу пропускну здатність та здатність адаптуватися до збільшення обсягів даних, що робить її перспективним рішенням для сучасної транспортної індустрії.

Література

1. Грицук В.Ю. Аналіз та обґрунтування вибору структури даних для завдань управління транспортними потоками. Вісник ХНАДУ, вип. 106, 2024. С.7-13. DOI: 10.30977/VUL.2219-5548.2024.106.0.7
2. Грицук В., Пронін С. Аналіз та обґрунтування вибору моделі для моніторингу параметрів транспортного засобу та прогнозування технічного обслуговування. Вісник Приазовського Державного Технічного Університету. Серія: Технічні науки, 1(49), 56–73. <https://doi.org/10.31498/2225-6733.49.1.2024.321206>
3. Волков В.П., Грицук І.В., Грицук Ю.В., Волков Ю.В., Володарець М.В. Інформаційні системи моніторингу технічного стану автомобілів. Харків: ФЛП Панов А.М., 2018. 299 с.
4. Neural network models // scikit-learn developers. – URL: https://scikit-learn.org/stable/modules/neural_networks_supervised.html (дата доступу 28.01.2025)
5. Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Parietal, INRIA Saclay Neurospin, Bat 145, CEA Saclay ^ 91191 Gif sur Yvette – France “Scikit-learn: Machine learning in python”, 2011.
6. Turk M. Eigenfaces for recognition / M. Turk, A. Pentland // Journal of Cognitive Neuroscience. – 1991. – Vol. 13, No. 1. – P. 71–86.
7. Shahadi, H. I., Haider, H. T., & Muhsen, D. H. (2018). Vehicle monitoring based on GSM technology for safety and security. IOP Conference Series: Materials Science and Engineering, 433(1). <https://doi.org/10.1088/1757-899X/433/1/012089> (дата доступу 28.01.2025)