https://doi.org/10.31891/2219-9365-2025-81-52
УДК 629.083

МАТІЙЧУК Любомир
Тернопільський національний технічний університет імені Івана Пулюя
https://orcid.org/0000-0001-6701-4683
e-mail: mlpstat@gmail.com
ГОТОВИЧ Володимир
Тернопільський національний технічний університет імені Івана Пулюя
https://orcid.org/0000-0003-2143-6818
e-mail: gotovych@gmail.com
БОНАР Віталій
Тернопільський національний технічний університет імені Івана Пулюя
https://orcid.org/0009-0002-6601-2584
e-mail: bonarv96@gmail.com

ПОРІВНЯННЯ ЕФЕКТИВНОСТІ МЕТОДІВ НЕКЕРОВАНОГО МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ В OBD2 ДАНИХ

У статті розглядається проведений експеримент у знаходженні аномальних станів двигуна базуючись на різних сигналах датчиків транспортного засобу (швидкість автомобіля, температура охолоджуючої рідини двигуна, температура моторного масла, кількість обертів двигуна у хвилину, абсолютне положення педалі акселератора, навантаження на двигун, витрата палива). Дані зібрані за допомогою адаптеру OBD2 зі справного дизельного автомобіля Honda CR-V. Отримані дані відфільтровано, систематизовано та нормалізовано перед початком обробки і знаходження аномалій.

У статті розглядається 7 методів некерованого машинного навчання знаходження аномалій - метод ізоляційного лісу, однокласовий метод опорних векторів, метод помилки реконструкції автоенкодера, метод аналізу головних компонент, метод дистанції К–середніх, коефіцієнт локального відхилення для знаходження аномалій, модель Гауссової суміші, метод глибокого опорного векторного опису даних.

Ключові слова: некероване машинне навчання, віддалена діагностика, OBD2, діагностичні коди несправностей.

MATIICHUK Liubomyr, HOTOVYCH Volodymyr, BONAR Vitalii
Ternopil Ivan Puluj National Technical University

COMPARISON OF THE EFFICIENCY OF UNSUPERVISED MACHINE LEARNING METHODS FOR DETECTING ANOMALIES IN OBD2 DATA

This article delves into an experimental investigation aimed at the identification of anomalous engine conditions through the analysis of diverse signals obtained from vehicle sensors. The study meticulously examines data streams originating from a properly functioning diesel Honda CR-V, acquired via an OBD2 adapter. The specific sensor signals under scrutiny encompass crucial operational parameters such as vehicle speed, engine coolant temperature, motor oil temperature, engine revolutions per minute (RPM), absolute accelerator pedal position, engine load, and fuel consumption.

Prior to the application of anomaly detection algorithms, the collected raw data underwent a rigorous preprocessing pipeline. This involved filtering to remove noise and inconsistencies, systematic organization to structure the data effectively, and normalization techniques to ensure that all features contribute equally to the subsequent analysis, mitigating the impact of differing scales and ranges.

The core of this research lies in the comparative evaluation of seven distinct unsupervised machine learning methodologies for the task of anomaly detection in the context of engine health monitoring. The methods explored include: the Isolation Forest algorithm, known for its efficiency in isolating outliers; the One-Class Support Vector Machine (OCSVM), adept at defining a boundary around normal data; the Autoencoder Reconstruction Error method, which identifies anomalies based on deviations in the reconstructed data; the Principal Component Analysis (PCA) method, leveraging dimensionality reduction to highlight deviations from the principal components; the K-means Distance method, which flags data points far from cluster centroids; the Local Outlier Factor (LOF) coefficient method, identifying anomalies based on their local density compared to neighbors; the Gaussian Mixture Model (GMM), which models the data as a mixture of Gaussian distributions and identifies low-probability points; and the Deep Support Vector Data Description (Deep SVDD) method, a deep learning approach for learning a compact hypersphere around normal data.

The study aims to provide a comprehensive comparative analysis of the performance of these unsupervised learning techniques in detecting abnormal engine states based on real-world vehicle sensor data. The findings of this research hold significant potential for the development of proactive vehicle maintenance systems, enabling early detection of potential engine malfunctions and contributing to enhanced vehicle safety and reliability. The comparative insights gained from evaluating these diverse methodologies will offer valuable guidance for selecting the most appropriate anomaly detection approach for automotive diagnostics.

Keywords: unsupervised machine learning, remote diagnostics, OBD2, diagnostic trouble codes.

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Швидкий розвиток автомобільних технологій призвів до інтеграції складних електронних систем керування в сучасні транспортні засоби. Сучасні транспортні засоби дозволяють отримати доступ до сигналів різних датчиків за допомогою порту OBD2. Отримані дані можуть бути використані для раннього

превентивного виявлення несправностей, що має вирішальне значення для підвищення безпеки, надійності та ефективності обслуговування автомобіля.

Традиційні методи діагностики на основі статистики та різних правил значною мірою покладаються на попередньо визначені порогові значення та експертні знання, які можуть бути обмежені у виявленні тонких або раніше невидимих аномалій. На противагу, методи неконтрольованого машинного навчання пропонують багатообіцяючу альтернативу - знаходження аномалій у нових невідомих даних, вивчаючи нормальну поведінку системи безпосередньо з тренованих даних. Це дозволяє виявляти відхилення, не вимагаючи прикладів помилок.

Збір даних для експерименту був проведений на справному дизельному автомобілі Honda CR-V. Для збору даних було використано Android додаток Car Scanner, який, за допомогою технології Bluetooth, підключається до OBD2 адаптеру vLinker FS Model CV304. Дана схема дозволяє використовувати всі можливості OBD2 протоколу для запиту необхідних даних автомобіля за допомогою ідентифікатора (PID) сигналу. Зібрані дані можна експортувати у необхідному форматі (в нашому випадку CSV) для їх подальшого аналізу.

Особливість збору даних за допомогою OBD2 полягає в тому, що читання даних відбувається по сигнально та ініціюється читачем [1] - в нашому випадку Android додатком. Експортований CSV файл має наступні дані (рис. 1):

1) SECONDS - часова відмітка запиту в секундах;
2) PID - ідентифікатор сигналу. У кінцевому файлі вже локалізовано у формат, зрозумілий людині;
3) VALUE - значення сигналу;
4) UNITS - міра вимірку сигналу;
5) LATITUDE та LONGITUDE - геопозиція автомобіля в даний момент часу, що додана Android додатком.

```
"SECONDS";"PID";"VALUE";"UNITS";"LATITUDE";"LONGTITUDE";
"26061.396161";"Calculated engine load value";"24.3137254901961";"%";"49.7683161062459";"24.021585633434";
"26061.524161";"Engine coolant temperature";"62";"°C";"49.7683161062459";"24.021585633434";
"26061.598161";"Calculated boost";"0.01";"bar";"49.7683161062459";"24.021585633434";
"26061.598161";"Intake manifold absolute pressure";"99";"kPa";"49.7683161062459";"24.021585633434";
"26061.683161";"Engine RPM";"820";"rpm";"49.7683161062459";"24.021585633434";
"26061.683161";"Engine RPM x1000";"0.8";"rpm";"49.7683161062459";"24.021585633434";
"26061.799161";"Vehicle acceleration";"0";"m/s²";"49.7683161062459";"24.021585633434";
"26061.799161";"Vehicle speed";"0";"km/h";"49.7683161062459";"24.021585633434";
```

Рис. 1. Приклад CSV файлу

В даному експерименті було поставлено питання “як двигун реагує на дії водія та умови навантаження, включаючи температурний стан та споживання палива?”. Для цього можна виділити наступні сигнали:

1) Швидкість автомобіля (Vehicle speed) - вимірюється у км/год;
2) Температура охолоджуючої рідини двигуна (Engine coolant temperature) - вимірюється у ℃;
3) Температура моторного масла (Engine oil temperature) - вимірюється у ℃;
4) Кількість обертів двигуна у хвилину (Engine RPM);
5) Абсолютне положення педалі D (Absolute pedal position D) - абсолютне положення педалі акселератора у %;
6) Абсолютне положення педалі E (Absolute pedal position E) - дублюючий сигнал, необхідний для виявлення помилок, якщо сигнал D несправний;
7) Розраховане значення навантаження двигуна (Calculated engine load value) - навантаження на двигун у даний момент часу у %;
8) Розрахована миттєва витрата палива (Calculated instant fuel rate) - витрата палива у даний момент часу у л/год.

Для ефективного аналізу даних потрібно отримати зріз необхідного стану автомобіля у певний момент часу, тому дані потрібно додатково попередньо обробити та нормалізувати. Перш за все, дані були відфільтровані, щоб включати лише сигнали, що перелічені вище.

Наступний крок - це нормалізація даних, а саме - створення зрізу всіх сигналів з інтервалом у певний проміжок часу, значення якого, шляхом експерементування і тестування було встановлено в одну секунду. Складність полягає в тому, що деякі сигнали можуть бути відсутні для певного зрізу. Для цього можна використати різні стратегії заповнення відсутніх даних. В даному експерименті було використано методи

заповнення вперед (forward fill) та назад (backward fill). Кінцева структура даних для аналізу представлена на рис. 2. Усі методи некерованого машинного навчання в даній статті використовують нормалізовані дані.

SECONDS	Absolute pedal position D	Absolute pedal position E	Calculated engine load value	Calculated instant fuel rate	Engine RPM	Engine coolant temperature	Engine oil temperature	Vehicle speed
1.0	20.0	9.0	24.313725	0.607376	820.0	62.0	57.0	0.0
2.0	20.0	9.0	24.313725	0.607376	820.0	62.0	57.0	0.0
3.0	20.0	9.0	24.313725	0.607376	820.0	62.0	57.0	0.0
4.0	0.0	9.0	25.490196	0.607376	820.0	62.0	57.0	0.0
5.0	0.0	9.0	24.313725	0.607376	820.0	62.0	57.0	0.0
...	...	...	...	...	...	...	...	...
15593.0	10.0	10.0	24.901961	0.579700	820.0	89.0	89.0	0.0
15594.0	10.0	10.0	24.705882	0.579700	820.0	89.0	89.0	0.0
15595.0	10.0	10.0	24.313725	0.597217	821.0	89.0	89.0	0.0
15596.0	10.0	10.0	24.313725	0.583088	821.0	89.0	89.0	0.0
15597.0	10.0	10.0	24.313725	0.583088	821.0	89.0	89.0	0.0

Рис. 2. Нормалізовані дані

Для ефективного тестування моделей некерованого машинного навчання окрім навчального масиву даних, потрібно створити і тестовий масив даних. Тестовий масив даних було створено на базі навчального з додаванням шуму за допомогою алгоритму Гауссівського шуму [3].

Метод ізоляційного лісу (Isolation Forest) [4] — це неконтрольований метод машинного навчання для виявлення аномалій, який базується на ідеї, що аномалій «небагато та вони різні», і тому їх легше виділити від решти даних. Перевага методу полягає в тому, що навчальні дані можуть містити аномалії.

Замість моделювання нормального розподілу даних, як у кластеризації, ізоляційний ліс працює шляхом випадкового розділення даних. Кожен розділ створюється шляхом вибору випадкової функції та випадкового значення розділення в діапазоні цієї функції. Цей процес продовжується рекурсивно, створюючи бінарне дерево, доки певна точка даних не буде ізольована. Кількість поділів, необхідних для ізоляції точки, називається довжиною шляху. Точки, які є аномаліями, як правило, швидко виділяються — іншими словами, вони мають меншу довжину шляху, оскільки розташовані далеко від щільних областей даних.

Для визначення, чи дані аномальні, визначають оцінку аномалії (anomaly score). Для кожної точки x, обчислюють середню довжину шляху $h(x)$ у всіх деревах. Це значення порівнюється із середньою довжиною шляху $c(n)$ в бінарному дереві пошуку такого ж розміру, де n - кількість даних. Оцінка аномалії $s(x) = 2^{-\frac{h(x)}{c(n)}}$ порівнюється з певним значенням для визначення аномальних даних.

Модель була навчена за допомогою методу Ізоляційного лісу на навчальному наборі даних і протестована на навчальному і тестовому наборах. Експериментально граничне значення оцінки аномалії було експериментально встановлено -0.1. На рис. 3 зображено результати моделі на навчальному (зліва) і тестовому (справа) наборах. На графіку видно, що аномалії було знайдено на обох наборах даних. Це свідчить про наявність хибно позитивної помилки.

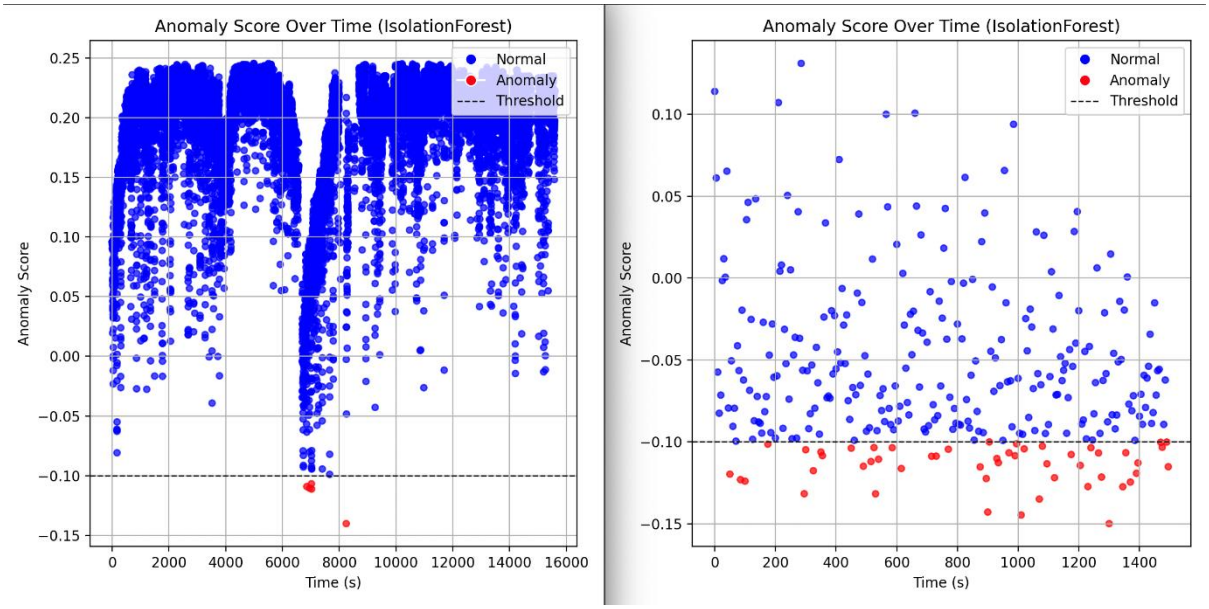

Рис. 4. Метод ізоляційного лісу (Isolation Forest)

Однокласовий метод опорних векторів (One-Class SVM) [5] - це неконтрольований метод виявлення аномалій, який описує область у просторі, де знаходяться не аномальні дані. Ідея полягає в тому, щоб знайти функцію, яка охоплює більшість точок даних у просторі, не включаючи точки, які знаходяться далеко від центру. Метод передбачає навчання на наперед відомих не аномальних даних.

Для опису алгоритму часто використовують поняття гіперсфери. В цьому випадку задача зводиться до знаходження мінімального значення радіусу гіперсфери, яка охопить усі не аномальні дані $R^2 \geq |x_i - a|^2$, де x_i - певна точка, a - центер гіперсфери.

Граничне значення оцінки аномалії було обрано -1. На рис. 5 видно, що, на відміну від методу Ізоляційного лісу, хибно позитивні помилки відсутні на навчальному наборі даних.

Знаходження аномалій за допомогою Автоенкодера (Autoencoder) [6] базується на аналізі помилки реконструкції після процесу кодування-декодування. Автокодер складається з двох основних компонентів: кодера, який перетворює вхідні дані в маловимірний латентний простір, і декодера, який реконструює першопочатковий вхідний сигнал із закодованого латентного простору.

Для вхідного вектора даних x, енкодер застосовує не лінійну функцію трасформації, яка створює кодоване значення z. Декодер, у свою чергу, намагається реконструювати значення x з z, в результаті отримуючи x'. Процес кодування-декодування містить помилку реконструкції, яка обчислюється за формулою $S = (x - x')^2$.

В контексті знаходження аномалій, передбачається, що енкодер буде навчений на заздалегідь відомих не аномальних даних для мінімізації помилки реконструкції. Відповідно, для не аномальних даних помилка реконструкції буде меншою, для аномальних - більшою.

На рис. 6 зображено результати знаходження аномалій за допомогою автоенкодера. Граничне значення помилки реконструкції було експериментально обрано 5. Даний метод показав досить високі результати точності знаходження аномалій.

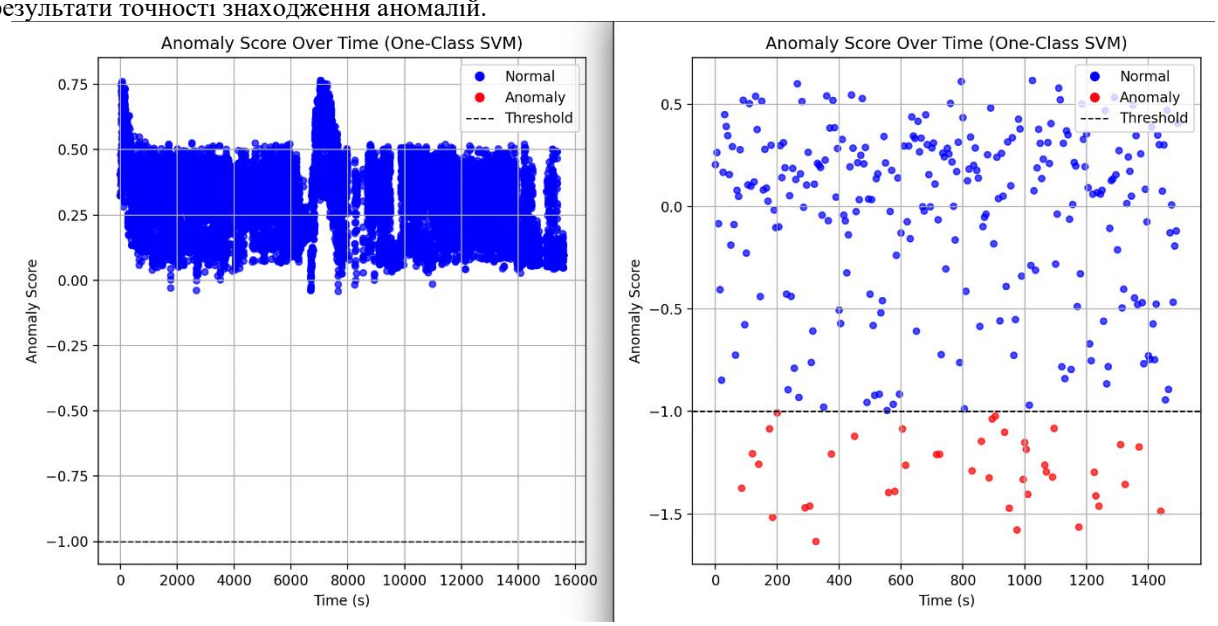

Рис. 5. Однокласовий метод опорних векторів (One-Class SVM)

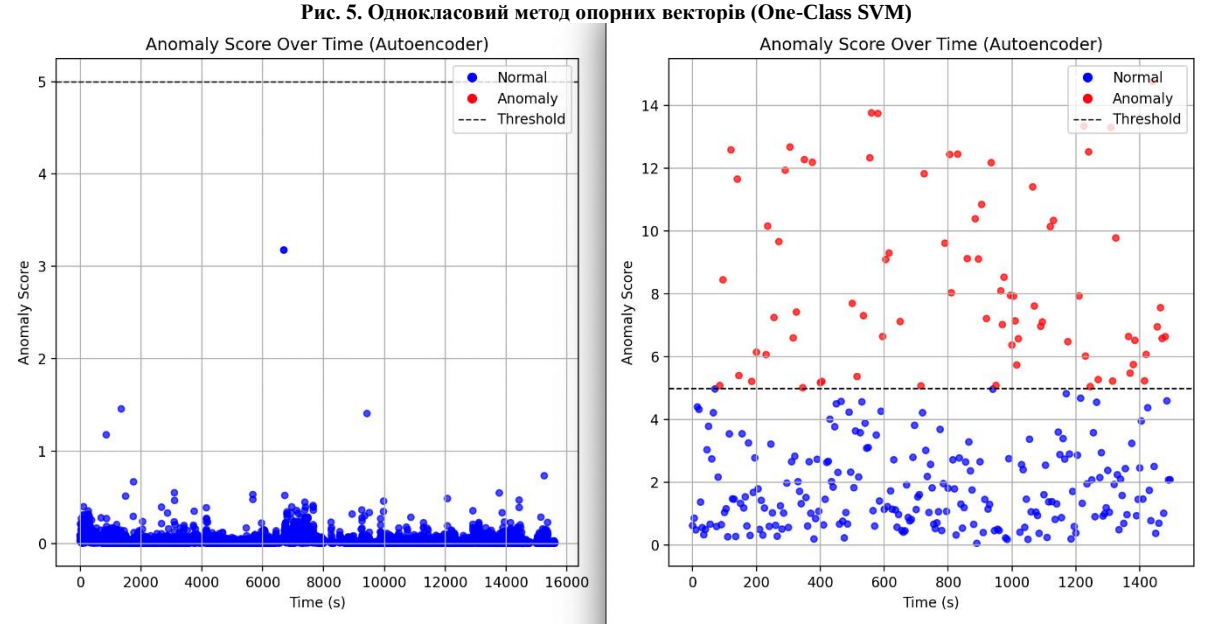

Рис. 6. Знаходження аномалій за допомогою Автоенкодера (Autoencoder)

Метод аналізу головних компонент (PCA) з помилкою реконструкції для знаходження аномалій базується на припущенні, що нормальні дані лежать близько до маловимірного лінійного підпростору у багатовимірному просторі. Аномалії, навпаки, це точки, які значно відхиляються від цього підпростору.

Завдання аналізу головних компонент полягає в тому, щоб зменшити розмірність деяких багатовимірних точок даних шляхом лінійного проектування їх на маловимірний простір таким чином, щоб помилка реконструкції, зроблена цією проєкцією, була мінімальною.

Як у випадку з автоенкодером, потрібно обрати певне граничне значення помилки реконструкції. Значення 5 для похибки було експериментально визначено. На рис. 7 зображено результати визначення аномалій методом головних компонент. Ефективність та точність подібна до методу автоенкодера.

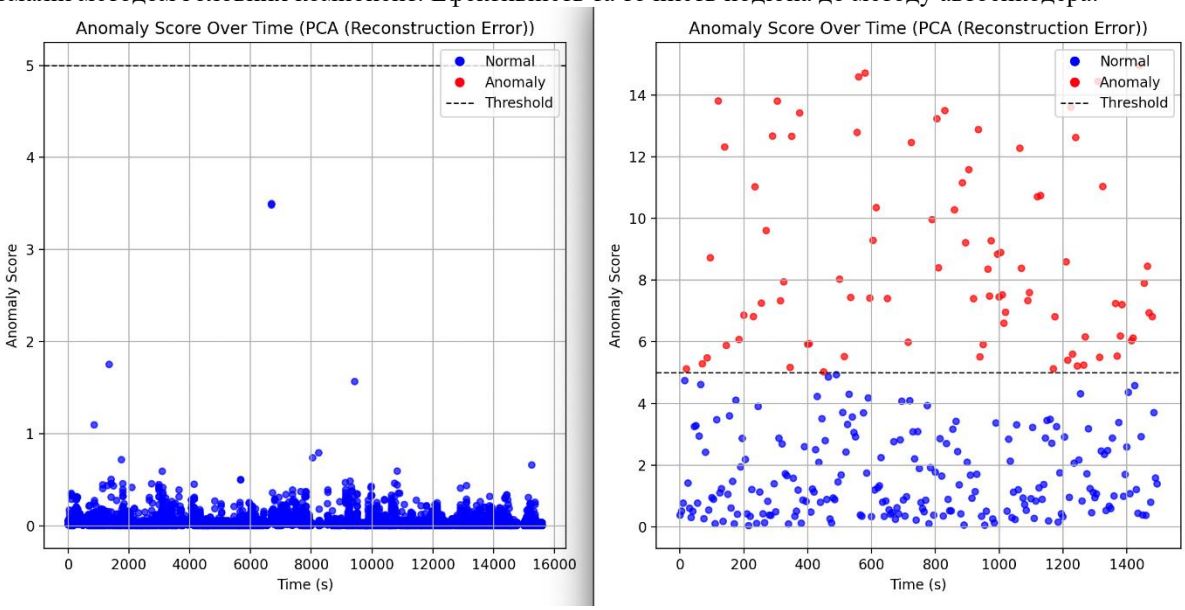

Рис. 7. Метод аналізу головних компонент (PCA)

Метод дистанції K–середніх (K-means distance) [8] - метод, який базується на ідеї, що звичайні точки даних мають тенденцію групуватися разом, тоді як аномалії знаходяться далеко від цих кластерів. Цей підхід використовує структуру кластеризації, отриману за допомогою алгоритму K-середніх, щоб обчислити оцінку аномалії як відстань точки від її найближчого центру кластера.

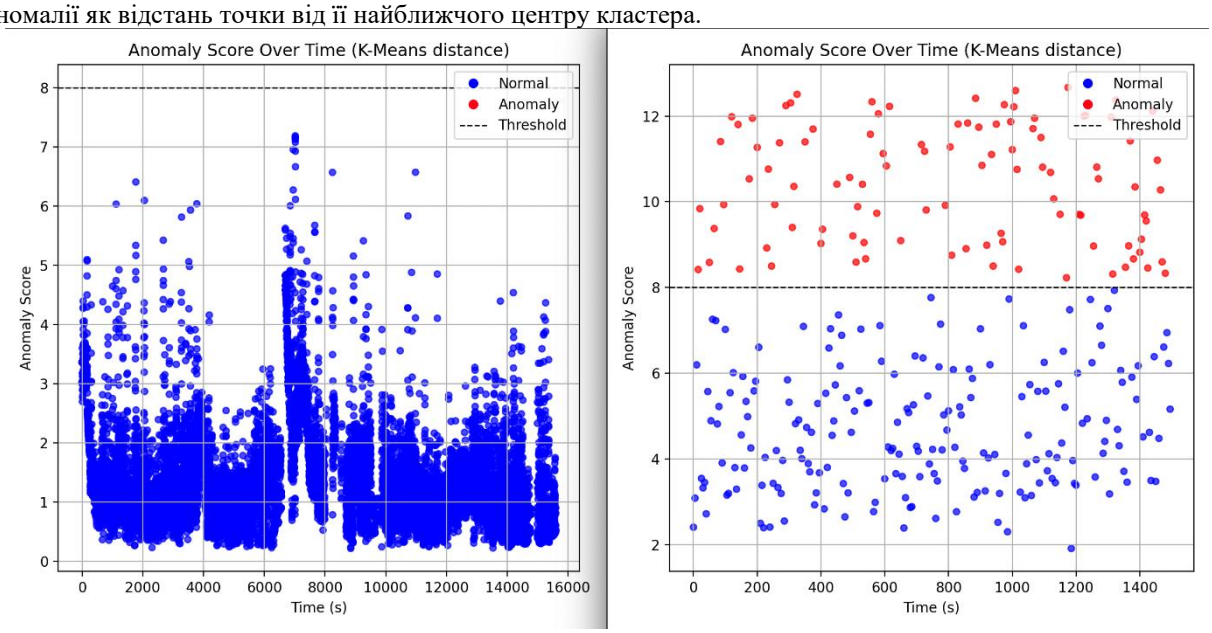

Рис. 8. Метод дистанції K–середніх (K-means distance)

Дані розділяються на K кластерів, мінімізуючи суму квадратів Евклідових відстаней між кожною точкою та центроїдом кластера, до якого вона належить. Математично задача зводиться до знаходження мінімальної відстані від кластера до точки $S = \min(x_i - u_k)^2$, де x_i - точка, u_k - кластер. Після процесу кластеризації обчислюється оцінка аномалії для кожної точки $score(x) = \min(x - u_k)$.

Метод дистанції K–середніх простий, швидкий та інтуїтивно зрозумілий. Для цього методу було експериментально обрано граничне значення оцінки аномалії 8. На рис. 8 відображено результати, які показують надійність та ефективність даного алгоритму

Коефіцієнт локального відхилення для знаходження аномалій (Local Outlier Factor) [9] - це неконтрольований метод виявлення аномалій, який ідентифікує точки як аномалії, якщо вони мають значно меншу локальну щільність, ніж їхні сусіди. Він заснований на ідеї, що нормальні точки оточені іншими точками з такою ж щільністю, тоді як аномалії ізольовані або лежать у розріджених областях.

Перевага метода полягає в можливості усунення обмеження методів, що базовані на глобальній відстані, які можуть не працювати в наборах даних із регіонами різної щільності. Коефіцієнт локального відхилення забезпечує відносну міру локального відхилення шляхом порівняння локальної щільності точки із щільністю її сусідів. Загальна формула для знаходження коефіцієнта для точки: $LOF(x) = \frac{1}{k}\sum_1^k \frac{density(y_k)}{density(x)}$, де x - задана точка, k - кількість сусідніх точок, y_k - сусідня точка. Коефіцієнт локального відхилення є оцінкою аномалії і порівнюється з визначеним граничним значенням для знаходження аномальних точок.

На рис. 9 зображено результати визначення аномалій даним методом з граничним значенням оцінки аномалії 3. Метод продемонстрував найбільш кількість хибно позитивних результатів на навчальному наборі даних.

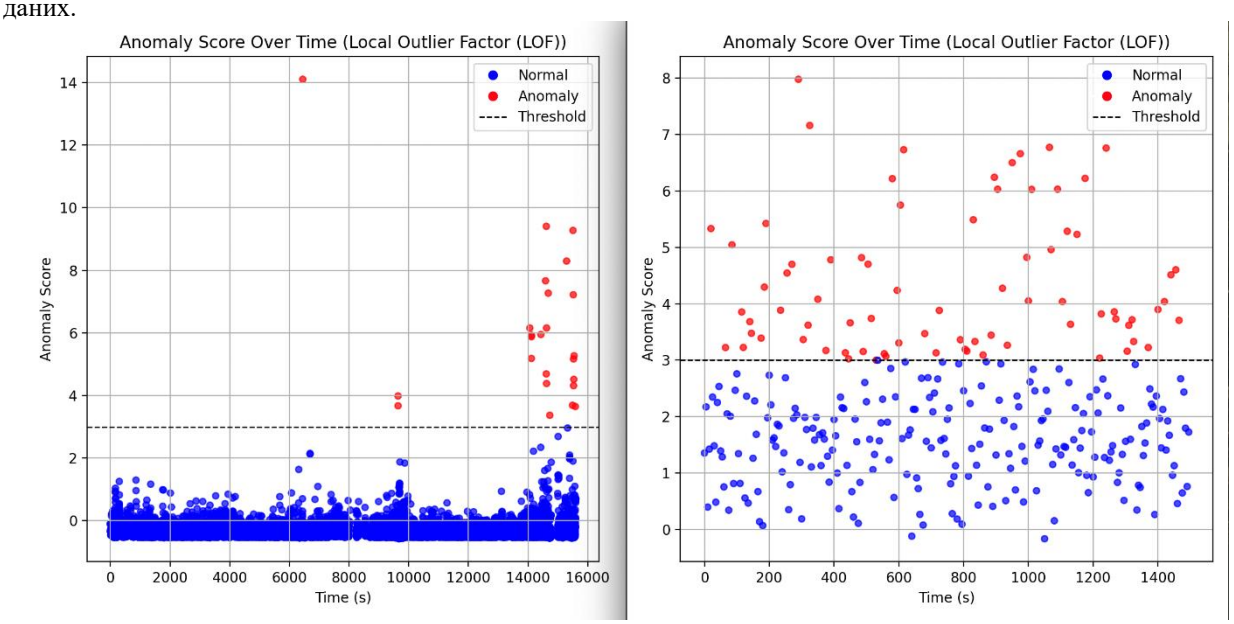

Рис. 9. Коефіцієнт локального відхилення для знаходження аномалій (Local Outlier Factor)

Модель Гауссової суміші (GMM) - це техніка м'якого кластеризування, яка використовується в неконтрольованому навчанні для визначення ймовірності, що дана точка даних належить кластеру. Модель складається з кількох нормальних розподілів Гаусса, кожен з яких ідентифікується, як $k \epsilon \{1, \ldots, K\}$, де K — кількість кластерів у наборі даних. Модель Гауссової суміші використовує ймовірнісний підхід для кластеризації даних, дозволяючи кластерам мати більш різноманітні форми та м'які межі.

Модель моделює щільність ймовірності даних як зважену суму Гауссівських розподілів $K: p(x) = \sum_{k=1}^{K} \pi_k \cdot N(x \mid u_k,\ \Sigma_k)$, де x - точка, π_k - зважена сума k -го Гауссівського розподілу, u_k - середній вектор k -го розподілу Гаусса, Σ_k- коваріаційна матриця k -го розподілу Гаусса, $N(x \mid u_k,\ \Sigma_k)$ - k -ий розподіл Гаусса. Оцінка аномальних даних обчислюється за формулою $score(x) = -\log p\,(x)$.

На рис. 10 зображено Модель Гауссової суміші з експериментально встановленою граничною оцінкою аномалії 450. Метод показав гарні результати на тестових та навчальних даних не маючи негативно позитивних результатів.

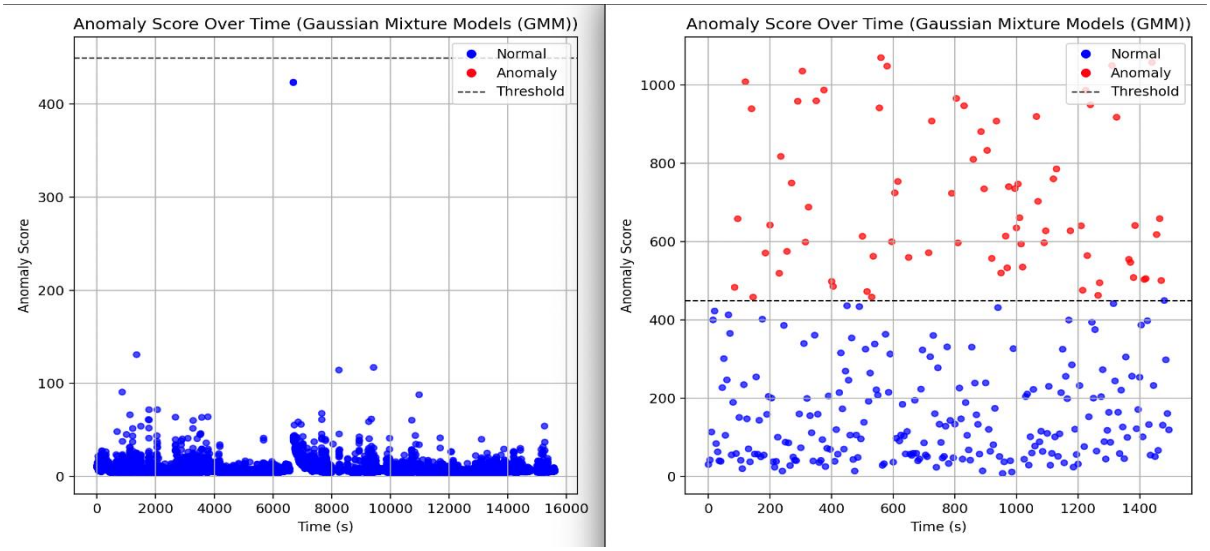

Рис. 10. Модель Гауссової суміші (GMM)

Метод глибокого опорного векторного опису даних (Deep SVDD (Support Vector Data Description)) [11] - це метод неконтрольованого машинного навчання для виявлення аномалій. Метод підходить для складних, багатовимірних даних.

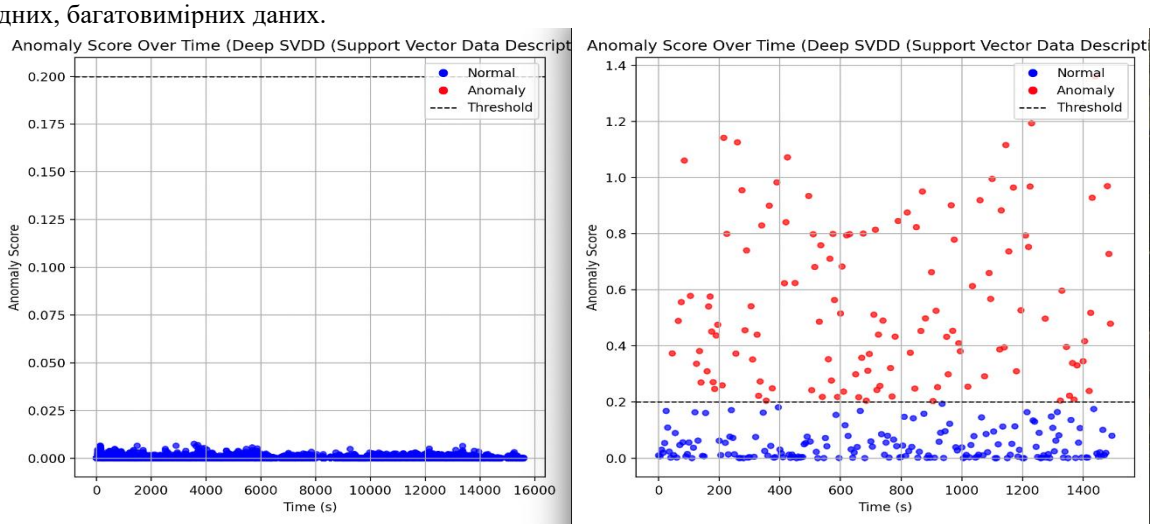

Рис. 11. Метод глибокого опорного векторного опису даних (Deep SVDD (Support Vector Data Description))

Основна ідея полягає в тому, щоб навчитися перетворювати вхідні дані в латентний простір, де розподіл не аномальних даних щільно згрупований навколо однієї точки, відомої як центр. Будь-яка точка даних, яка знаходиться значно далі від цього центру, вважається аномальною. На відміну від традиційного методу, який покладається на функції ядра для проектування даних у багатовимірний простір функцій, метод глибокого опорного векторного опису даних вивчає цю проекцію безпосередньо через нейронну мережу. Це дозволяє методу гнучко адаптуватися до складних, багатовимірних і нелінійних структур даних, таких як сигнали часових рядів від датчиків автомобіля.

Після навчання оцінка аномалії для будь-яких нових даних x обчислюється шляхом обчислення квадрата евклідової відстані між його латентним зображенням і центром: $score(x) = (f(x) - c)^2$, де $f(x)$ - функція нейронної мережі, c - центр гіперсфери.

На рис. 11 зображено результати знаходження аномалій методом глибокого опорного векторного опису даних з граничним значенням оцінки аномалії 0.2. Метод показав гарні результати знаходження аномальних точок, при цьому не додаючи негативно позитивних результатів.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПДАЛЬШОГО РОЗВИТКУ У ДАНОМУ НАПРЯМІ

Було проведено експеримент у знаходженні аномальних станів двигуна базуючись на сигналах (швидкість автомобіля, температура охолоджуючої рідини двигуна, температура моторного масла, кількість обертів двигуна у хвилину, абсолютне положення педалі акселератора, навантаження на двигун, витрата палива) зібраних за допомогою OBD2 з дизельного автомобіля Honda CR-V.

Аномальні стани знайдено у тестовому наборі даних за допомогою різних методів некерованого машинного навчання. У знаходженні аномальних станів дуже важливо уникати негативно позитивних результатів, тобто, якщо точка не аномальна, а вона визначається аномальною. Зважаючи на це, найточніші результати показали ті методи, в яких не аномальні дані були зображені найщільніше на графіках: однокласовий метод опорних векторів, помилка реконструкції автоенкодера, метод аналізу головних компонент, модель Гауссової суміші. Менш точні: метод ізоляційного лісу, метод дистанції K–середніх, коефіцієнт локального відхилення для знаходження аномалій. Найбільш точний метод з найщільнішим розподілом навчальних даних - метод глибокого опорного векторного опису даних.

У майбутніх дослідженнях планується огляд статистичних методів та методів керованого машинного навчання.

Література

1. Rybitskyi, Oleksandr & Golian, Vira & Golian, Nataliia & Dudar, Zoia & Kalynychenko, Olga & Nikitin, Dmytro. (2023). USING OBD-2 TECHNOLOGY FOR VEHICLE DIAGNOSTIC AND USING IT IN THE INFORMATION SYSTEM. Bulletin of National Technical University KhPI Series System Analysis Control and Information Technologies. 97-103. 10.20998/2079-0023.2023.01.15.
2. Moahmed, T.A., El Gayar, N., Atiya, A.F. (2014). Forward and Backward Forecasting Ensembles for the Estimation of Time Series Missing Data. In: El Gayar, N., Schwenker, F., Suen, C. (eds) Artificial Neural Networks in Pattern Recognition. ANNPR 2014. Lecture Notes in Computer Science(), vol 8774. Springer, Cham. https://doi.org/10.1007/978-3-319-11656-3_9.
3. Zhang, Xinhua. (2016). Gaussian Distribution. 10.1007/978-1-4899-7502-7_107-1.
4. Liu, Fei Tony & Ting, Kai & Zhou, Zhi-Hua. (2009). Isolation Forest. 413 - 422. 10.1109/ICDM.2008.17.
5. Hejazi, Maryamsadat & Singh, YashwantPrasad. (2013). One-class support vector machines approach to anomaly detection. Applied Artificial Intelligence. 27. 351-366. 10.1080/08839514.2013.785791.
6. Ganesh, Mukkesh & Kumar, Akshay & Pattabiraman, V. (2020). Autoencoder Based Network Anomaly Detection. 1-6. 10.1109/TEMSMET51618.2020.9557464.
7. Huang, Yingkun & Jin, Weidong & Yu, Zhibin & Li, Bing. (2021). A robust anomaly detection algorithm based on principal component analysis. Intelligent Data Analysis. 25. 249-263. 10.3233/IDA-195054.
8. Münz, G., Li, S., & Carle, G. (2007). Traffic Anomaly Detection Using K-Means Clustering.
9. Arya Adesh, Shobha G, Jyoti Shetty, Lili Xu, Local outlier factor for anomaly detection in HPCC systems, Journal of Parallel and Distributed Computing, Volume 192, 2024, 104923, ISSN 0743-7315, https://doi.org/10.1016/j.jpdc.2024.104923.
https://www.sciencedirect.com/science/article/pii/S074373152400087X
10. Li, Lishuai & Hansman, R. & Palacios, Rafael & Welsch, Roy. (2016). Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring. Transportation Research Part C: Emerging Technologies. 64. 45-57. 10.1016/j.trc.2016.01.007.
11. Liu, Chenyu & Gryllias, Konstantinos. (2021). A Deep Support Vector Data Description Method for Anomaly Detection in Helicopters. PHM Society European Conference. 6. 10.36001/phme.2021.v6i1.2957.

References

1. Rybitskyi, Oleksandr & Golian, Vira & Golian, Nataliia & Dudar, Zoia & Kalynychenko, Olga & Nikitin, Dmytro. (2023). USING OBD-2 TECHNOLOGY FOR VEHICLE DIAGNOSTIC AND USING IT IN THE INFORMATION SYSTEM. Bulletin of National Technical University KhPI Series System Analysis Control and Information Technologies. 97-103. 10.20998/2079-0023.2023.01.15.
2. Moahmed, T.A., El Gayar, N., Atiya, A.F. (2014). Forward and Backward Forecasting Ensembles for the Estimation of Time Series Missing Data. In: El Gayar, N., Schwenker, F., Suen, C. (eds) Artificial Neural Networks in Pattern Recognition. ANNPR 2014. Lecture Notes in Computer Science(), vol 8774. Springer, Cham. https://doi.org/10.1007/978-3-319-11656-3_9.
3. Zhang, Xinhua. (2016). Gaussian Distribution. 10.1007/978-1-4899-7502-7_107-1.
4. Liu, Fei Tony & Ting, Kai & Zhou, Zhi-Hua. (2009). Isolation Forest. 413 - 422. 10.1109/ICDM.2008.17.
5. Hejazi, Maryamsadat & Singh, YashwantPrasad. (2013). One-class support vector machines approach to anomaly detection. Applied Artificial Intelligence. 27. 351-366. 10.1080/08839514.2013.785791.
6. Ganesh, Mukkesh & Kumar, Akshay & Pattabiraman, V. (2020). Autoencoder Based Network Anomaly Detection. 1-6. 10.1109/TEMSMET51618.2020.9557464.
7. Huang, Yingkun & Jin, Weidong & Yu, Zhibin & Li, Bing. (2021). A robust anomaly detection algorithm based on principal component analysis. Intelligent Data Analysis. 25. 249-263. 10.3233/IDA-195054.
8. Münz, G., Li, S., & Carle, G. (2007). Traffic Anomaly Detection Using K-Means Clustering.
9. Arya Adesh, Shobha G, Jyoti Shetty, Lili Xu, Local outlier factor for anomaly detection in HPCC systems, Journal of Parallel and Distributed Computing, Volume 192, 2024, 104923, ISSN 0743-7315, https://doi.org/10.1016/j.jpdc.2024.104923. https://www.sciencedirect.com/science/article/pii/S074373152400087X
10. Li, Lishuai & Hansman, R. & Palacios, Rafael & Welsch, Roy. (2016). Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring. Transportation Research Part C: Emerging Technologies. 64. 45-57. 10.1016/j.trc.2016.01.007.
11. Liu, Chenyu & Gryllias, Konstantinos. (2021). A Deep Support Vector Data Description Method for Anomaly Detection in Helicopters. PHM Society European Conference. 6. 10.36001/phme.2021.v6i1.2957.