

МОДЕЛЮВАННЯ ВІДМОВ ПРОМИСЛОВИХ ЕЛЕКТРОДВИГУНІВ З ВИКОРИСТАННЯМ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ

Н. А. Заєць, доктор технічних наук, професор

Національний університет біоресурсів і природокористування України

Н. М. Луцька, доктор технічних наук, професор

Національний університет харчових технологій

Л. О. Власенко, кандидат технічних наук, доцент

Національний університет біоресурсів і природокористування України

E-mail: z-n@ukr.net, vlasenko.lidia1@gmail.com, lutskanm2017@gmail.com

Анотація. Проведено аналіз, сформульовано та визначено класифікацію нештатних ситуацій, що виникають у результаті переходу із штатного до нештатного режиму функціонування електротехнологічного комплексу харчових виробництв. Розроблена структура системи підтримки прийняття рішень для моніторингу електротехнологічних процесів підприємства в реальному масштабі часу з метою локалізації прогнозованих або виявлених у результаті моніторингу нештатних ситуацій. Розроблено програмно-апаратний комплекс системи моніторингу та підтримки прийняття рішень, основним завданням якого є зменшення незапланованих зупинок виробництва та простою обладнання при виникненні нештатних ситуацій за рахунок прогнозування роботи електротехнологічного комплексу та збільшення швидкості реакції диспетчера при управлінні технологічним комплексом харчових виробництв.

Ключові слова: нештатна ситуація, аварія, система керування, підтримка прийняття рішень, харчове виробництво

Актуальність. Останнім часом штучний інтелект широко застосовується для діагностики та прогнозування появи несправностей та управління продуктивністю виробництва [1]. Використання алгоритмів машинного навчання активно застосовується для моніторингу стану промислового обладнання [2], зокрема промислових двигунів. Дані, що характеризують нормальну роботу двигуна, та дані, що свідчать про появу несправностей, можна отримати з традиційної системи автоматизації, зі спеціально введеної системи моніторингу стану двигунів та з журналів поломок. Традиційні системи автоматизації зазвичай можуть надати інформацію про аварійну зупинку двигуна, або проблеми із його зупинкою чи

частотою обертання. Журнали поломок є обов'язковою формою документування всіх несправностей промислового обладнання, включаючи електродвигуни, насоси тощо, ремонтною групою виробничого відділу. Побудова промислового підприємства відповідно до принципів Industry 4.0 передбачає впровадження інтелектуальних систем для обробки даних та побудови відповідних моделей [3]. Такі системи можуть автоматично аналізувати характеристики вхідних даних для прогнозування появи виникнення несправностей промислового обладнання [4], здійснювати ефективне планування технічного обслуговування, що є критичним для забезпечення ефективності роботи промислового обладнання, зокрема, двигунів. Можливість впровадження технологій IoT [5] та хмарних технологій дозволяє використовувати складніші методи прогнозування аварій при мінімальних затратах часу та ресурсів.

Об'єкт дослідження – моделювання поломок та надійності промислових електродвигунів для покращення системи моніторингу та технічного обслуговування.

Предмет дослідження - модель поломок промислових електродвигунів на основі даних отриманих з традиційний системи автоматичного керування САК та журналів поломок ремонтних бригад промислового підприємства.

Аналіз останніх досліджень та публікацій. Ефективне виявлення несправностей важливе у будь-якій галузі обслуговування, оскільки своєчасне розпізнавання проблем може запобігти появі серйозних відмов і простоїв, сприяти економії часу і фінансових ресурсів. Звичайні методи виявлення поломок зазвичай потребують великих витрат, включаючи як часові, так і людські ресурси. Це сприяло зростанню інтересу до розробки методів машинного навчання, зокрема, алгоритмів неконтрольованого навчання, з метою оптимізації процесів діагностики.

Визначення несправностей, засноване на принципах машинного навчання, є загальноприйнятий підхід, що використовує дані для автоматичного вивчення діагностичних знань без прив'язки до експертного досвіду чи інженерних знань. Цей метод ґрунтується на теорії машинного навчання для автоматизованого вивчення

зв'язків між зібраними даними та станом обладнання для створення відповідної моделі.

Для ефективного функціонування виробництва необхідно забезпечити надійну і безперебійну роботу електродвигунів [6]. Надійність електродвигуна оцінюється його спроможністю безперебійно працювати протягом значного періоду часу за різних умов експлуатації, враховуючи стійкість до різноманітних впливів, таких як механічні навантаження, температурні зміни, вологість, напругові коливання та інші фактори. Для оцінки надійності враховують: середній наробіток на відмову (MTBF – Mean Time Between Failures), ймовірність відмови, ступінь надійності, резервуар для усунення збоїв.

Найпоширеніші методи машинного навчання, які застосовуються для аналізу та прогнозування стану двигунів можна розділити на три групи [7, 8]:

- Поверхневе машинне навчання (Shallow machine learning – SML). Моделі поломок на основі SML являють собою процес визначення закономірностей, що передбачає застосування методів обробки сигналів для виокремлення ознак несправності та формування векторів характеристик. Для завершення ідентифікації несправностей використовуються класифікатори, такі як штучні нейронні мережі (Artificial neural network – ANN) та машини опорних векторів (Support vector machine – SVM) [9].

- Глибоке навчання (Deep learning – DL). Моделі поломок на основі DL містять більшу кількість шарів у нейронних мережах, націлені на повне використання великих об'ємів даних для отримання точніших результатів [10].

- Трансферне навчання (Transfer learning – TL). Трансферне навчання полягає в тому, щоб передати наявні знання для вирішення проблем навчання в цільовій області з невеликою кількістю позначених зразків або навіть без них. Більше того, чим більше факторів поділяються на два різні домени, тим простіше провести навчання [11].

Модель сумішей Гауса (Gaussian Mixture Model – GMM) – це функція ймовірності, що має параметри і представлена як вагова сума густин гаусівських компонентів. Зазвичай GMM використовується як параметрична модель для

розподілу ймовірностей неперервних вимірювань або характеристик у технічних системах, наприклад, у визначенні особливостей роботи двигунів для позначення появи аномалій під час їх функціонування. Параметри GMM визначаються через ітераційний алгоритм очікування-максимізації (ЕМ) або максимального апостеріорного (MAP) оцінювання на основі даних, отриманих добре навченою моделі.

Існують різні підходи використання GMM для моніторингу появи поломок двигунів. У [12] представлені результати розробки моделі кластеризації, ідентифікації та діагностики шести різних несправностей в електродвигунах на основі поєднання GMM та відстані Махаланобіса. Існує низка робіт, що підтверджують ефективність використання GMM для покращення надійності і точності оцінки, виявлення та ідентифікації несправностей підшипників двигунів [13, 14].

Взагалі спектр застосування методів машинного навчання досить широкий: це і виявлення короткого замикання в синхронних двигунах [15], і різні види поломок для асинхронних двигунів відцентрових насосів [16], і систем багатофазного приводу [17], і виявлення аномалій в асинхронних двигунах з інверторним живленням [18] тощо.

Мета дослідження - розробка моделей поломок промислових електродвигунів з використанням методів навчання з учителем (unsupervised learning) з метою ефективного виявлення аномалій та прогнозування можливих несправностей.

Матеріали і методи дослідження. При моделюванні надійності електродвигунів на основі поточних вимірювань їх характеристик з використанням моделей машинного навчання скористаймося алгоритмом, представленим на рис. 1.

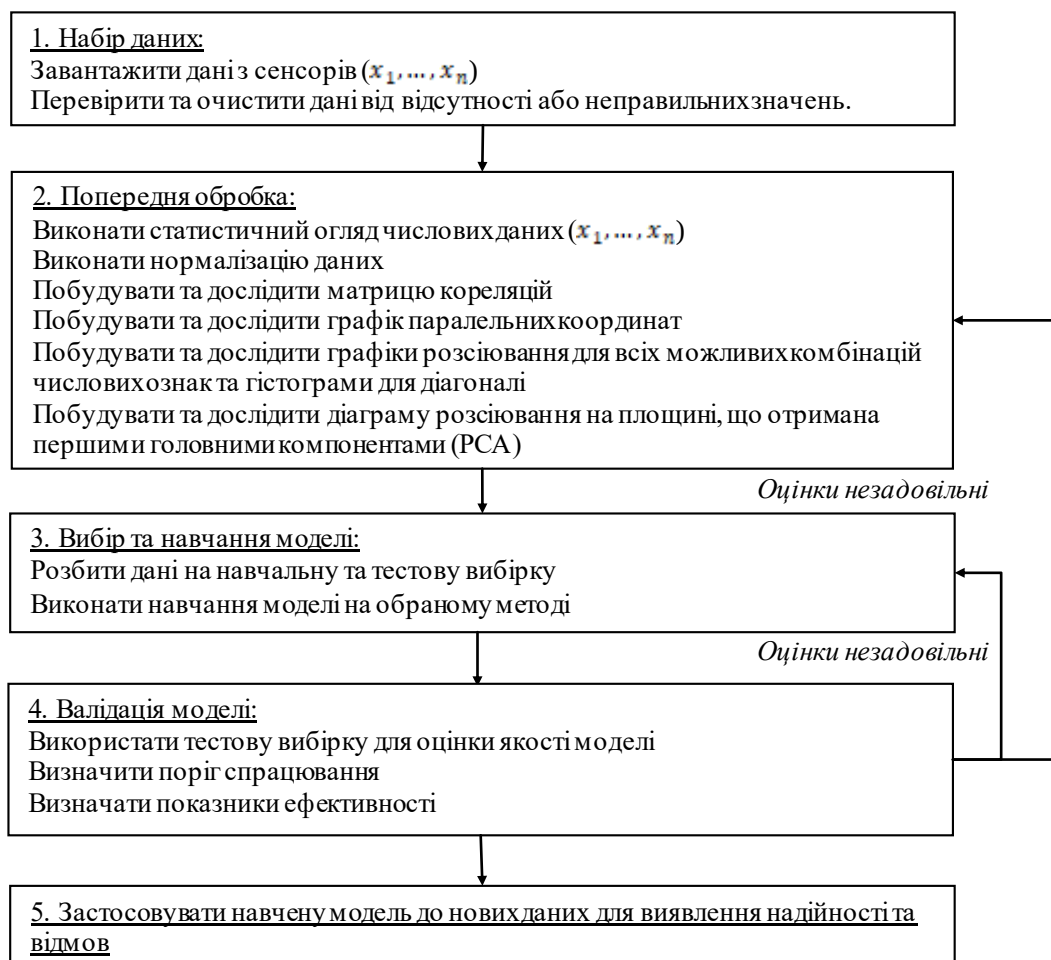


Рис. 1. Алгоритм процесу моделювання несправності електродвигунів

Першим етапом є отримання даних, що репрезентативно відображають надійність електродвигунів, такі як температура, витрати струму, напруга, частота обертання, тощо. На різних виробництвах харчової промисловості можуть вимірюватися різні ознаки, що можуть відображати надійність двигуна.

Найчастіше до таких ознак належать:

- середнє навантаження (Average load, % – x_1);
- частота обертання (Rotational frequency, rpm – x_2);
- напруга живлення (Power supply voltage, V – x_3);
- споживаний струм (Current consumption, A – x_4);
- струм витікання (Leakage current, mA – x_5).

Особливістю набору даних є велика кількість даних, які позначені як надійні та мала кількість даних, які позначені як несправні. Кількість останніх, як правило, становить <15 % від загальної кількості. Крім того, враховуючи різнотипність

поломок двигунів та малої кількості даних для кожного типу поломки, ефективним буде використання методів виявлення аномалій. Такі методи використовуються для виявлення несподіваних або надто відхиляючихся від патернів даних для нормальної роботи системи.

Після отримання набору даних, їхньої перевірки та попереднього очищення потрібно їх розділити на надійні дані та дані поломок та розглядати окремо кожен набір.

Другим етапом є перегляд даних та їх перетворення. На цьому етапі необхідно виконати статистичний огляд числових даних для кожного розподілу та виконати нормалізацію даних. Остання забезпечить стабільність та швидку збіжність алгоритму навчання моделі та уніфікую вплив ознак на модель.

Дослідження матриці кореляцій допоможе підтвердити вплив ознак на вихідний показник, однак ознаки з малим коефіцієнтом кореляції не слід відкидати. Низький коефіцієнт кореляції не обов'язково вказує на відсутність взаємозв'язку або на те, що ознака не має впливу на вихідний показник. Також може виникнути ситуація, коли взаємозв'язок є, але він не виявляється за допомогою лінійної кореляції (яка вимірює тільки лінійні взаємозв'язки). Ознаки з малим коефіцієнтом кореляції також можуть мати важливий вплив на модель, особливо якщо вони взаємодіють з іншими ознаками або виявляють свій вплив не за допомогою лінійного зв'язку. Тому важливо ретельно аналізувати всі можливі взаємозв'язки та брати до уваги всі аспекти впливу ознак на вихідний показник. Якщо ознака вважається важливою з точки зору контексту задачі (наприклад, наведені вище), навіть при відсутності сильної кореляції, необхідно ретельніше досліджувати її вплив або розглядати інші методи виявлення впливу ознак на вихідний показник.

Дослідження графіка паралельних координат дозволяє візуалізувати об'єкти в просторі багатовимірних ознак, де кожна лінія представляє один об'єкт, а координатні вісі представляють різні ознаки. Діаграма розсіювання на площині, що отримана першими головними компонентами (РСА), надає візуальне представлення структури даних у власній системі координат, де вони максимально змінюються. Разом ці дві візуалізації допоможуть виявити загальну структуру даних та розподіл

спостережень у високорозмірних даних, забезпечуючи спрощену інтерпретацію та візуалізацію.

Побудова графіків розсіювання для всіх можливих комбінацій числових ознак та їх гістограм дозволить зручно візуалізувати залежності між парами змінних у наборі даних, а також допоможе зрозуміти взаємозв'язків у багатовимірних даних. Крім того, працюючи з гістограмами можна визначитися з подальшими методами побудови моделей, зокрема обрати метрику, що найкраще описує дані. Також допоможе зрозуміти чи обраний алгоритм нормалізації вірний і чи потрібне додаткове перетворення ознак.

Наступним етапом є вибір та навчання моделі на навчальній вибірці. Враховуючи особливості даних, які зазначалися раніше, навчальна вибірка повинна містити лише дані, які відносяться до надійної роботи двигунів. Тестовий набір повинен містити як надійні дані обладнання, так і їх поломки. Навчання моделі виконується відповідно до всіх рекомендацій ML.

Після навчання моделі необхідно ретельно перевірити її на тестових даних та підібрати поріг спрацювання відмови електродвигуна. Для цього можна застосувати такі показники ефективності:

- Ассурасу – загальний показник того, наскільки правильно модель класифікує нормальні та аномальні (поломки) зразки. Рекомендується використовувати лише разом з іншими.

- Recall або True Positive Rate – вимірює здатність моделі виявляти аномалії серед всіх дійсно аномальних випадків або здатність моделі правильно класифікувати нормальні випадки. Рекомендується досягти високих значень обох показників.

- F1-score є збалансованою мірою між Recall та Precision. Вона обчислюється як гармонічне середнє між ними. Враховуючи нерівномірний розподіл класів корисно використовувати F1- score для класу поломок, щоб оцінити ефективність моделі щодо їх виявлення.

- AUC-ROC – вимірює здатність моделі розділяти класи, а саме, ймовірність того, що аномальний зразок отримає вищу оцінку, ніж нормальний.

- Likelihood Ratio – особливо корисний для визначення границі прийняття аномалії. Для даної задачі висока вірогідність може слугувати показником високої надійності електродвигуна.

При незадовільних оцінках ефективності моделі потрібно повернутися до попередніх пунктів та переглянути обрану метрику отримання параметрів моделі та/або попередню обробку даних.

Отриману модель можна застосовувати для оптимізації роботи електродвигунів, підвищення їх надійності та зменшення ризиків виникнення аварійних ситуацій.

Результати досліджень та їх обговорення. Виконаємо процес моделювання надійності електродвигунів потужністю 55 кВт, 1500 об/хв, дані про які отримано на цукровому заводі. Процес моделювання виконаємо за допомогою мови програмування Python, з використанням бібліотек pandas, matplotlib, seaborn та sklearn. В якості середовища програмування обрано Google Colaboratory.

Особливістю цього набору даних є велика кількість даних з високою надійністю ($y = 1$) та мала кількість з низькою ($y = 0$) – 2857/42. На рис. 2, *a* зображено перші п'ять точок даних та основні статистичні характеристики отриманих ознак. Ймовірно дані мають значну варіативність, що пов'язана з даними, які ідентифікують поломки. Можна зробити висновок про необхідність подальшого нормування.

Розглянемо матрицю кореляцій та графік паралельних координат для нормованих значень ознак (рис. 3). Всі ознаки, окрім першої мають високий кореляційний коефіцієнт стосовно вихідної змінної, також x_2 та x_3 мають високий лінійний зв'язок, а x_4 також корелюється з x_5 . Однак на графіку паралельних координат для надійних пристроїв ($y = 1$) можна побачити подібність, що підтверджує існування єдиного кластеру нормальної роботи двигунів. І навпаки, при несправній роботі ($y = 0$) ми спостерігаємо декілька кластерів (як мінімум два), однак висновок щодо кількості кластерів при аномаліях зробити не можемо, тому що даних занадто мало.

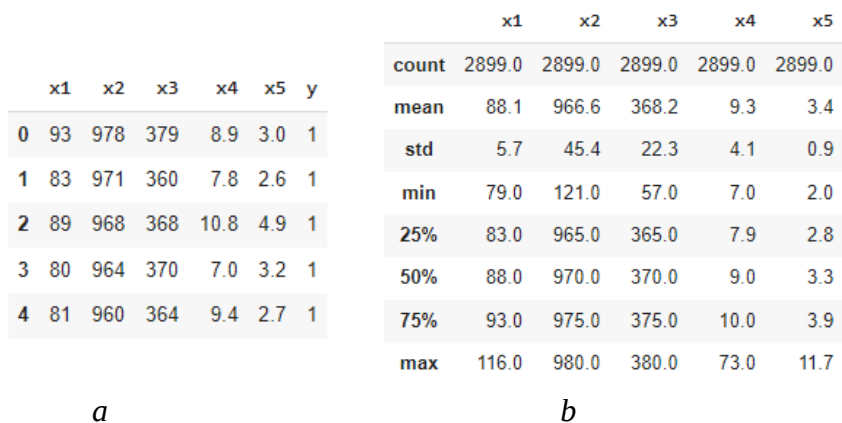


Рис. 2. Перші значення в наборі даних (*a*) та статистичні оцінки ознак (*b*)

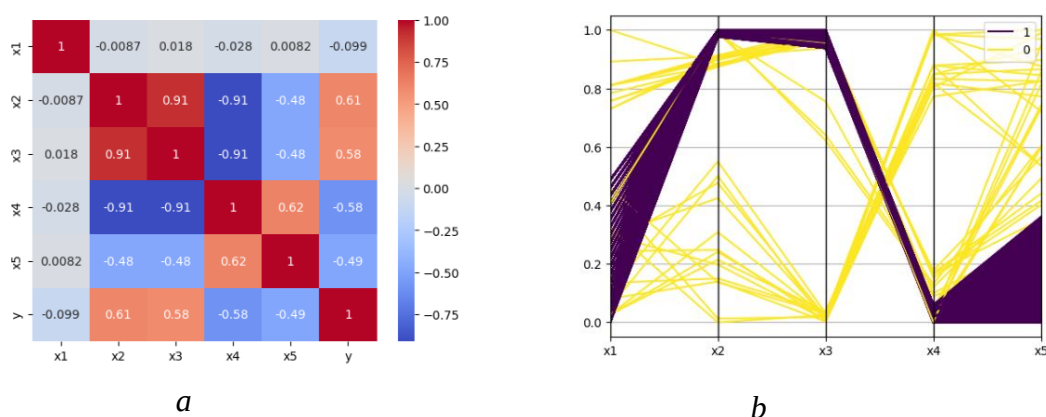


Рис. 3. Теплова карта даних (*a*) графік паралельних координат для нормованих значень ознак (*b*)

Після побудови графіка попарного розсіювання для даних, що характеризуються надійною роботою (рис. 4, *a*), можна спостерігати на міждіагональних графіках характерні патерни, а на гістограмах (діагональні графіки) можна припустити, що розподіл ймовірності кожної ознаки апроксимується сумішшю Гауса з кількістю компонент 4 або 5. Таким чином, можна обрати метод для моделювання надійності електродвигунів – GMM, що є параметричною моделлю та базується на параметрах розподілу суміші Гауса. Рис. 4, *b* також нам демонструє ймовірність існування двох кластерів для даних, що описують поломки двигунів.

На рис. 5 зображено діаграму розсіювання, що побудована на площині двох головних компонент (Principal Component 1 та 2). Можна спостерігати основні напрямки варіації в даних. Нормальні об'єкти схильні групуватися разом, тоді як

аномалії вирізняються як окремі точки, які ймовірно утворюють три кластера. Один кластер досить близько підходить до нормальних точок даних.

Після побудови моделі GMM, що моделює надійність електроприводів отримали такі параметри:

$$p(x) = \sum_{i=1}^{K=4} \pi_i \mathcal{N}(x; \mu_i; \sigma_i), \quad (1)$$

де K – кількість компонент у суміші гаусіан ($K = 4$); π_i – вага i -го компонента (ймовірність вибору i -го компонента):

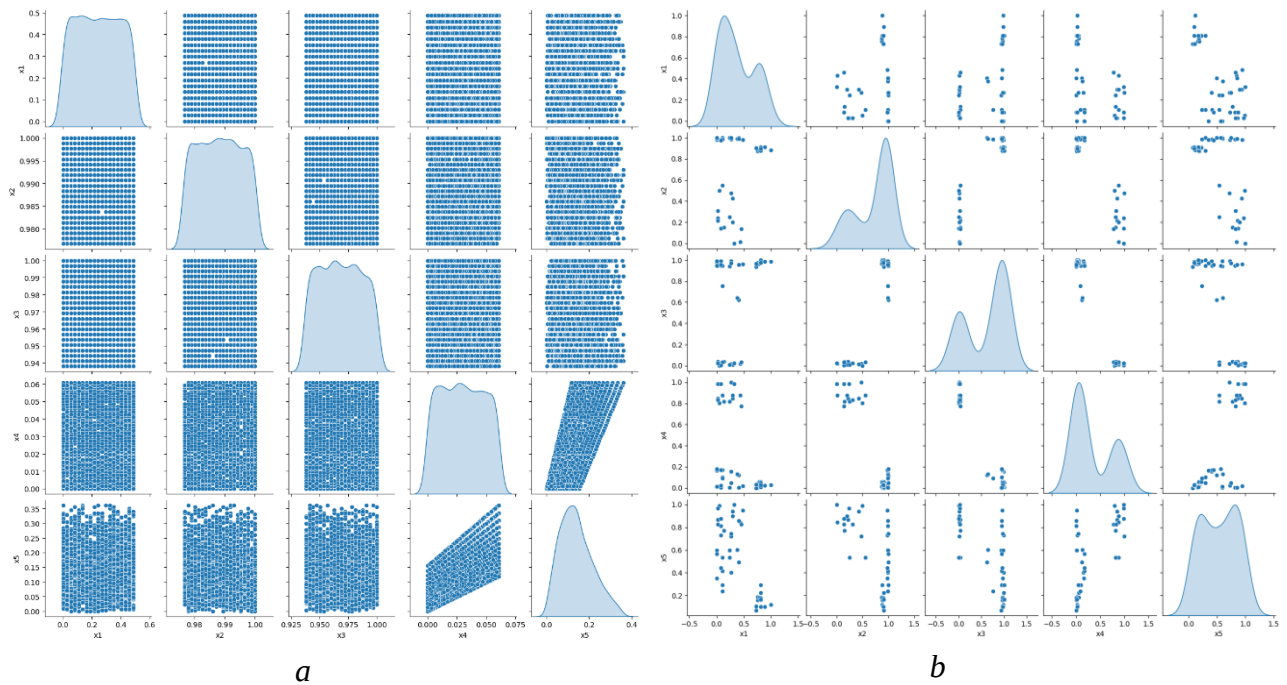


Рис. 4. Графік парних взаємозв'язків у наборі даних для $y = 1$ (a) та $y = 0$ (b)

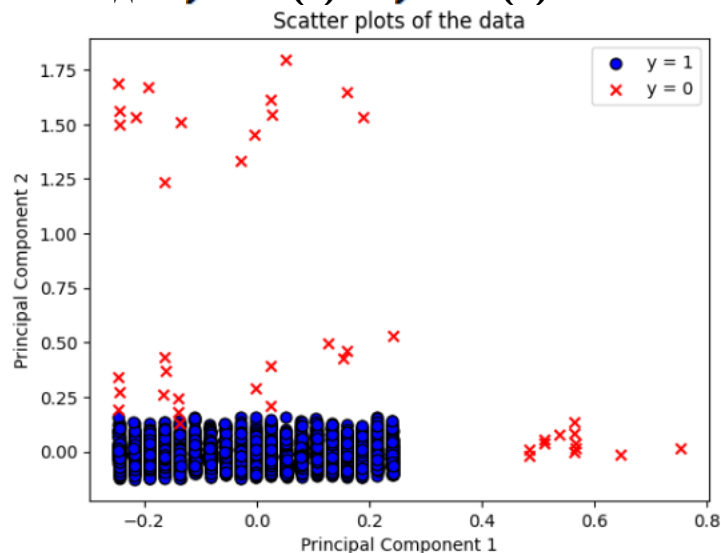


Рис. 5. Точковий графік даних навчання на площині головних компонент

$$\pi_{i=1...K} = [0.39; 0.28; 0.16; 0.17]; \quad (2)$$

$\mathcal{N}(x; \mu_i; \sigma_i)$ – функція Гауса щільності ймовірності для i -го компонента з параметрами μ_i (середнє значення):

$$\mu_{i=1...K} = \begin{bmatrix} 0.21 & 0.99 & 0.97 & 0.03 & 0.09 \\ 0.27 & 0.99 & 0.97 & 0.03 & 0.20 \\ 0.05 & 0.99 & 0.97 & 0.03 & 0.16 \\ 0.44 & 0.99 & 0.97 & 0.03 & 0.13 \end{bmatrix}; \quad (3)$$

та σ_i (коваріаційна матриця) відповідних розмірностей для багатовимірного випадку (наведено лише першу компоненту):

$$\sigma_{i=1} = \begin{bmatrix} 1.26e-02 & 1.80e-05 & -6.25e-05 & -5.15e-04 & -9.02e-04 \\ 1.80e-05 & 4.96e-05 & -2.39e-06 & 3.45e-06 & 1.61e-05 \\ -6.25e-05 & -2.39e-06 & 3.42e-04 & 2.12e-05 & 2.37e-05 \\ 5.15e-04 & 3.45e-06 & 2.12e-05 & 3.32e-04 & 7.01e-04 \\ -9.02e-04 & 1.61e-05 & 2.37e-05 & 7.01e-04 & 2.15e-03 \end{bmatrix}, \quad (4)$$

Для перевірки отриманої моделі скористаймося графіком розсіювання для тестових даних (Рис. 6). Поріг імовірності виявлення аномальної поведінки електродвигунів був підбраний так, щоб всі поломки були ідентифіковані GMM. Таким чином, отримуємо наступну вираз для прийняття рішень щодо поломок обладнання:

$$\text{if } p(x) < \text{threshold then } y = 0, \text{ threshold} = 5.0 \quad (5)$$

Як видно з графіку розсіювання (рис. 6) отримано 100 % підгонку даних на тестовому наборі. Це саме доводить аналіз показників ефективності, який наведений в табл. 1. Зокрема, False Positive Rate (FPR) = 0.0, тобто відсутні хибно позитивні випадки – усі об'єкти, які насправді не є позитивними, правильно класифіковані як негативні. Точність на тренувальних даних не досягає 1.0 і це є нормальним, тому що ми намагалися уникнути перенавчання та забезпечити узагальнення моделі на нових даних.

1. Метрики оцінювання

Метрика	Навчальні дані	Тестувальні дані
Precision	0.99	1.0
Recall	1.0	1.0
F1-score	0.99	1.0
Accuracy	0.99	1.0
AUC-ROC	0.99	1.0
False Positive Rate	0.0	0.0

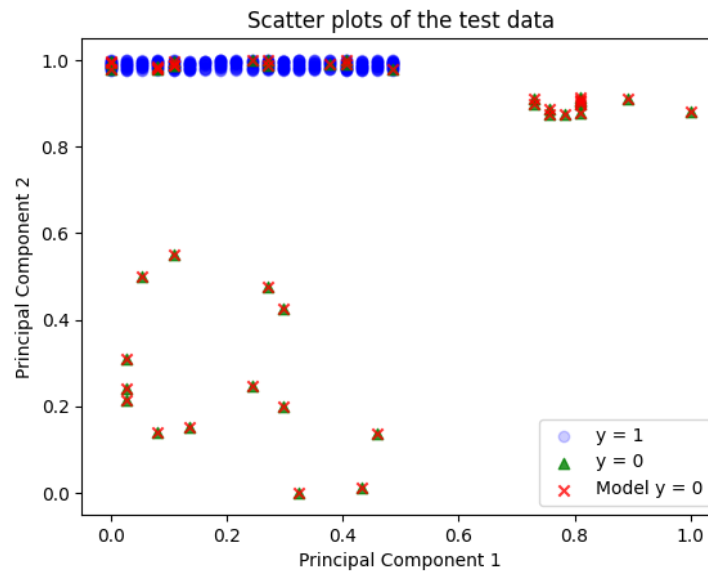


Рис. 6. Точковий графік даних тестування на площині головних компонент

Розглянемо графік аномалій на шкалі логарифмічної ймовірності (рис. 7). Знову спостерігаємо два кластера, що виявляють аномалію. Один зосереджений в області низької логарифмічної ймовірності, інший – ближче до кластера нормальної роботи двигунів. Імовірно існує третій кластер низької надійності об'єктів, який перетинається з кластером $y = 0$. Тобто, існують дві локальні аномалії (невеликі кластери низької ймовірності) та одна глобальна аномалія (велика область низької ймовірності). Це все вказує на різні типи аномалій у даних, що потребує додаткове вивчення природи аномальних кластерів та причин їх появи.

Із аналізу рис. 7 можна візуально вказати поріг спрацювання аномалій. Однак, автори скористалися автоматичним інструментом для оцінки ефективності систем виявлення аномалій, що оцінює компроміс між помилками детекції (detection error tradeoff – DET), який знайшов поріг $threshold = 6.47$, що несуттєво відрізняється від отриманого порогу вище.

Таким чином, використовуючи отриману модель можна прогнозувати несправності двигунів, а подальший аналіз отриманих патернів допоможе виявити причини несправності електроприводів.

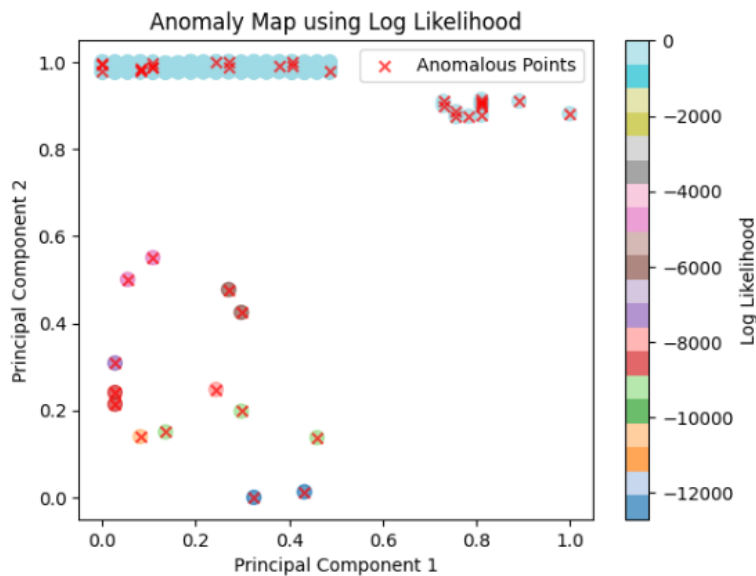


Рис. 7. Карта аномалій з використанням логарифмічної правдоподібності тестових даних на площині головних компонент

Висновки і перспективи. Модель виявлення надійності та поломок електродвигунів, побудована за алгоритмом виявлення аномалій з використанням методів машинного навчання, може бути використана для ряду завдань та застосувань в області управління та технічного обслуговування, зокрема:

1. Виробничий моніторинг обладнання та виявлення аномалій – модель допоможе вчасно виявляти аномалії та невідповідності, що може вказувати на можливі поломки чи проблеми з роботою обладнання, уникне аварійних ситуацій.
2. Планування обслуговування та ремонтів – виявлені на основі моделі аномалії допоможуть спланувати роботи з технічного обслуговування та ремонту. Аналізуючи отримані патерни, модель може надавати інформацію про те, які частини чи компоненти електродвигунів можуть потребувати уваги або заміни.
3. Оптимізація витрати енергії – модель можна використати для оптимізації витрати енергії та підвищення ефективності.
4. Вдосконалення роботи технічних систем – своєчасно виявлені на основі моделі аномалії можна використати для постійного вдосконалення технічних систем, враховуючи знайдені проблеми при розробці нових моделей чи стратегій управління.

5. Взаємодія із системами управління – можна побудувати зворотний зв'язок управління при виявленні аномалій, що враховує резервні зв'язки, пристрої тощо.

6. Візуалізація та звітність – модель може стати основою для інтерфейсу візуалізації результатів роботи електроприводів та створення звітів для операторів або інших зацікавлених сторін.

У цілому застосування цього підходу та розробленої моделі допоможе оптимізувати роботу електродвигунів, підвищити їх надійність та зменшити ризик виникнення аварійних ситуацій.

References

1. Vlasenko, L. O., Lutska, N. M., Zaiets, N. A., Shyshak, A. V., Savchuk, O. V. (2022). Domain ontology development for condition monitoring system of industrial control equipment and devices. *Radio Electronics, Computer Science, Control*, 1, 157. <https://doi.org/10.15588/1607-3274-2022-1-16>
2. Lutska, N., Vlasenko, L., Zaiets, N., Lysenko, V. (2022). Modeling the Productivity of a Sugar Factory using Machine Learning Methods. In *2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)*, 353-356. <https://doi.org/10.1109/CSIT56902.2022.10000571>
3. Korobiichuk, I., Mel'nick, V., Kosova, V., Ostapenko, Z., Gnateiko, N., Rzeplinska-Rykala, K. (2023). Mathematical Model of the Approximate Function as the Result of Identification of the Object of Automatic Control. In: *Conference on Automation*. Cham: Springer Nature Switzerland, 2023. 173-182. <https://doi.org/10.1108/IR-12-2021-0296>
4. Drakaki, M., Karnavas, Y. L., Tzifettas, I. A., Linardos, V., Tzionas, P. (2022). Machine learning and deep learning based methods toward industry 4.0 predictive maintenance in induction motors: State of the art survey. *Journal of Industrial Engineering and Management (JIEM)*, 15.1: 31-57. <https://doi.org/10.1109/TIA.2022.9757372>
5. Tran, M. Q., Amer, M., Abdelaziz, A. Y., Dai, H. J., Liu, M. K., Elsis, M. TRAN, Minh-Quang, et al. (2023). Robust fault recognition and correction scheme for induction motors using an effective IoT with deep learning approach. *Measurement*, 207: 112398. <https://doi.org/10.1016/j.measurement.2022.112398>
6. Korobiichuk, I., Kachniarz, M., Bezvesilna, O., Nowicki, M., Ilchenko, A., Szewczyk, R. (2017). Calorimetric flow meter of motor fuel With Inlet temperature regulation. In: *2017 4th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2017. 0975-0979. <https://doi.org/10.1109/CoDIT.2017.8102725>
7. Sunal, C. E., Dyo, V., Velisavljevic, V. (2022). NReview of machine learning based fault detection for centrifugal pump induction motors. *IEEE Access*, 2022, 10: 71344-71355. <https://doi.org/10.1109/ACCESS.2022.3187718>

8. Cen, J., Yang, Z., Liu, X., Xiong, J., Chen, H. (2022). A review of data-driven machinery fault diagnosis using machine learning algorithms. *Journal of Vibration Engineering & Technologies*, 2022, 10.7: 2481-2507. <https://doi.org/10.1007/s42417-022-00498-9>
9. Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., Nandi, A. K. LEI, Yaguo, et al. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 2020, 138: 106587. <https://doi.org/10.1016/j.ymssp.2019.106587>
10. Mushtaq, S., Islam, M. M., & Sohaib, M.(2021). Deep learning aided data-driven fault diagnosis of rotatory machine: A comprehensive review. *Energies*, 2021, 14.16: 5150. <https://doi.org/10.3390/en14165150>
11. Li, C., Zhang, S., Qin, Y., Estupinan, E. (2020). A systematic review of deep transfer learning for machinery fault diagnosis. *Neurocomputing*, 2020, 407: 121-135. <https://doi.org/10.1016/j.neucom.2020.04.045>
12. Ribeiro Junior, R.F., de Almeida, F.A., Jorge, A.B. et al. (2023). On the use of the Gaussian mixture model and the Mahalanobis distance for fault diagnosis in dynamic components of electric motors. *J Braz. Soc. Mech. Sci. Eng.* 45, 139. <https://doi.org/10.1007>
13. Lin, Shih-Lin. (2021)/ Application of machine learning to a medium Gaussian support vector machine in the diagnosis of motor bearing faults. *Electronics* 10.18: 2266. <https://doi.org/10.3390/electronics10182266>
14. Toma, Rafia Nishat, and Jong-myon Kim. (2021). Bearing Fault Classification of Induction Motor Using Statistical Features and Machine Learning Algorithms. *International Conference on Intelligent Systems Design and Applications*. Cham: Springer International Publishing, 243-254. https://doi.org/10.1007/978-3-030-96308-8_22
15. Shih, Kai-Jung, et al. (2022). Machine learning for inter-turn short-circuit fault diagnosis in permanent magnet synchronous motors." *IEEE Transactions on Magnetics* 58.8: 1-7. <https://doi.org/10.1109/TMAG.2022.3169173>
16. Sunal, Cem Ekin, Vladimir Dyo, and Vladan Velisavljevic. "Review of machine learning based fault detection for centrifugal pump induction motors." *IEEE Access* 10 (2022): 71344-71355. <https://doi.org/10.1109/ACCESS.2022.3187718>
17. Liu, Zicheng, et al. (2022). A machine-learning-based fault diagnosis method with adaptive secondary sampling for multiphase drive systems." *IEEE transactions on power electronics* 37.8: 8767-8772. <https://doi.org/10.1109/TPEL.2022/3153797>
18. Cherif, Bilal Djamal Eddine, et al. (2022). Machine-learning-based diagnosis of an inverter-fed induction motor." *IEEE Latin America Transactions* 20.6, 901-911. <https://doi.org/10.1109/TLA.2022.9757372>
19. Real-time expert system a real gold mine. Available at: <http://www.controlglobal.com/articles/2005/412>

MODELING INDUSTRIAL ELECTRIC MOTOR FAILURES USING MACHINE LEARNING ALGORITHMS

N. Zaiets, N. Lutska, L. Vlasenko

Abstract. *The task of detecting breakdowns helps to improve the reliability of industrial devices through early identification of problems and their timely resolution, which helps to improve the resource and energy efficiency of production as a whole. When modeling breakdowns of industrial electric motors, the following characteristics are primarily taken into account: average load, rotational frequency, power supply voltage, current consumption, leakage current, collected from data from a traditional automated control system and breakdown logs of repair teams of an industrial enterprise. The model for detecting reliability and breakdowns of electric motors in the work is built on an anomaly detection algorithm using unsupervised machine learning methods. Its effectiveness has been successfully tested using many evaluation metrics. It can be used for a range of tasks and applications in equipment management and maintenance, visualization, forecasting and management decision support.*

Key words: *unsupervised learning, Gaussian Mixture Model, motor, modelling, breakdown, reliability*