

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНЖЕНЕРІЇ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ**

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Методика оновлення обладнання підприємств зв'язку
з використанням прогнозної аналітики»

на здобуття освітнього ступеня магістра
зі спеціальності 121 Інженерія програмного забезпечення
освітньо-професійної програми «Інженерія програмного забезпечення»

*Кваліфікаційна робота містить результати власних досліджень. Використання
ідей, результатів і текстів інших авторів мають посилання
на відповідне джерело*

_____ Ольга ПЕРМЯКОВА
(підпис)

Виконала: здобувачка вищої освіти групи ПДМ-62

_____ Ольга ПЕРМЯКОВА

Керівник: _____ Світлана ШЕВЧЕНКО
канд. пед. наук, доц.

Рецензент: _____

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**
Навчально-науковий інститут інформаційних технологій

Кафедра Інженерії програмного забезпечення

Ступінь вищої освіти Магістр

Спеціальність 121 Інженерія програмного забезпечення

Освітньо-професійна програма «Інженерія програмного забезпечення»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Інженерії програмного забезпечення

_____ Ірина ЗАМРІЙ

« ____ » _____ 2025 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

_____ Пермяковій Ользі Олександрівні _____

1. Тема кваліфікаційної роботи: «Методика оновлення обладнання підприємств зв'язку з використанням прогнозної аналітики»

керівник кваліфікаційної роботи Світлана ШЕВЧЕНКО, канд. пед. наук, доц., доцент кафедри ІІЗ

затверджені наказом Державного університету інформаційно-комунікаційних технологій від «30» жовтня 2025 р. № 467.

2. Строк подання кваліфікаційної роботи «19» грудня 2025 р.

3. Вихідні дані до кваліфікаційної роботи: теоретичні відомості про телекомунікаційне обладнання, науково-технічна література, моделі штучного інтелекту, методи заміни телекомунікаційного обладнання.

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

1. Аналіз предметної галузі

2. Аналіз моделей штучного інтелекту для розробки методики

3. Розробка методу оновлення обладнання на підприємстві зв'язку з використанням прогнозної аналітики

4. Тестування системи

5. Перелік графічного матеріалу: *презентація*

1. Порівняльний аналіз існуючих методів оновлення обладнання.
2. Математична модель.
3. Алгоритм методики оновлення обладнання підприємств зв'язку.
4. Практичний результат.
5. Аналіз ефективності розробленої методики.
6. Апробація результатів дослідження.

6. Дата видачі завдання «31» жовтня 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз наявної науково-технічної літератури	31.10-05.11.2025	
2	Аналіз існуючих методів оновлення обладнання	06.11-11.11.2025	
3	Аналіз моделей штучного інтелекту для розробки методики	11.11-17.11.2025	
4	Розробка методу оновлення обладнання на підприємстві зв'язку з використанням прогнозної аналітики	18.11-30.11.2025	
5	Тестування системи	01.12-05.12.2025	
6	Оформлення роботи: вступ, висновки, реферат	06.12-15.12.2025	
7	Розробка демонстраційних матеріалів	16.12-19.12.2025	

Здобувачка вищої освіти

(підпис)

Ольга ПЕРМЯКОВА

Керівник
кваліфікаційної роботи

(підпис)

Світлана ШЕВЧЕНКО

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістра: 82 стор., 3 табл., 9 рис., 22 джерел.

Мета роботи – удосконалення процесу заміни та оновлення обладнання на підприємствах зв'язку за допомогою розробленої методики з використанням штучного інтелекту.

Об'єкт дослідження – процес оновлення обладнання на підприємствах зв'язку.

Предмет дослідження – методика оновлення обладнання підприємств зв'язку з використанням прогнозної аналітики.

У роботі використано різноманітні методи, такі як системний аналіз, математичне моделювання, методи машинного навчання.

Проведено аналіз сучасних методів управління життєвим циклом телекомунікаційного обладнання. Розроблено та удосконалено методику прогнозної аналітики. Створено математичну модель та розроблено алгоритм аналізу стану обладнання. Проведено тестування для оцінки ефективності розробленої методики, виконано оцінку часу виконання методики.

КЛЮЧОВІ СЛОВА: ТЕЛЕКОМУНІКАЦІЙНЕ ОБЛАДНАННЯ, ТЕХНІЧНЕ ОБСЛУГОВУВАННЯ, ПРОГНОЗНА АНАЛІТИКА, ШТУЧНИЙ ІНТЕЛЕКТ, МАШИННЕ НАВЧАННЯ.

ABSTRACT

Text part of the master's qualification work: 82 pages, 9 pictures, 3 table, 22 sources.

The purpose of the work is to improve the process of equipment replacement and renewal at telecommunication enterprises using a developed methodology based on artificial intelligence.

Object of research – the process of equipment renewal at telecommunication enterprises.

Subject of research – the methodology for equipment renewal at telecommunication enterprises using predictive analytics.

Summary of the work:

The thesis employs various methods, such as system analysis, mathematical modeling, and machine learning techniques.

An analysis of modern methods for managing the lifecycle of telecommunication equipment has been conducted.

A predictive analytics methodology has been developed and improved. A mathematical model has been created, and an algorithm for analyzing the equipment state has been developed.

Testing has been carried out to evaluate the effectiveness of the developed methodology, and the execution time of the methodology has been assessed.

KEYWORDS: TELECOMMUNICATION EQUIPMENT, MAINTENANCE, PREDICTIVE ANALYTICS, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING.

ЗМІСТ

ВСТУП.....	10
1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ	13
1.1 Огляд та аналіз літературних джерел.....	13
1.2 Аналіз процесу технічного обслуговування та заміни обладнання на підприємствах зв'язку.....	16
1.3 Класифікація телекомунікаційного обладнання та фактори впливу на його надійність	18
1.4 Аналіз існуючих рішень заміни обладнання на підприємствах зв'язку	21
2 АНАЛІЗ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ РОЗРОБКИ МЕТОДИКИ	25
2.1 Штучний інтелект у прогностному обслуговуванні (Predictive Maintenance). 25	
2.1.1 Основні переваги використання прогностної аналітики на підприємствах зв'язку	25
2.1.2 Оцінка вимог до даних та очікуваної точності моделей.....	27
2.2 Огляд моделей машинного навчання для прогнозування	29
2.3 Інтерпретація моделей та підтримка прийняття рішень	36
2.3.1 Методи пояснювального ШІ (XAI) для обґрунтування рекомендацій	36
2.3.2 Підходи до візуалізації прогнозів для інженерного персоналу	38
2.4 Проблеми обробки незбалансованих даних	39
3 РОЗРОБКА МЕТОДУ ОНОВЛЕННЯ ОБЛАДНАННЯ НА ПІДПРИЄМСТВІ ЗВ'ЯЗКУ З ВИКОРИСТАННЯМ ПРОГНОЗНОЇ АНАЛІТИКИ.....	42
3.1 Математична модель прогнозування оновлення обладнання	42
3.2 Нормалізація даних.....	44
3.3 Алгоритм для аналізу стану обладнання	46
3.4 Застосування математичної моделі	49
3.4.1 Засоби розробки та архітектура ПЗ.....	49
3.4.2 Програмна реалізація алгоритму.....	50
3.4.3 Візуалізація роботи алгоритму.....	55

4 ТЕСТУВАННЯ СИСТЕМИ	63
4.1 Оцінка часу виконання методики.....	63
4.2 Оцінка ефективності розробленої моделі	65
ВИСНОВКИ.....	67
ПЕРЕЛІК ПОСИЛАНЬ	69
ДОДАТОК А. ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ	72
ДОДАТОК Б. ЛІСТИНГ ПРОГРАМНИХ МОДУЛІВ	77

ВСТУП

У сучасних умовах цифрової трансформації підприємств зв'язку питання ефективного технічного обслуговування та своєчасної заміни обладнання набуває особливої важливості. Зношення апаратного забезпечення, зростання обсягів переданої інформації та підвищення вимог до надійності мереж спричиняють необхідність переходу до інтелектуальних методів управління життєвим циклом обладнання.

Традиційні стратегії обслуговування – реактивна та планово-попереджувальна – вже не відповідають цим вимогам. Реактивний підхід призводить до неконтрольованих аварійних простоїв, а планове обслуговування спричиняє значні фінансові втрати через надлишкову заміну повністю справного обладнання. Таким чином, актуальною задачею є не лише прогнозування відмови, а й забезпечення максимальної оптимізації CapEx та OpEx шляхом переходу до обслуговування Just-in-Time.

Застосування штучного інтелекту для прогнозування технічного стану систем, планування заміни компонентів і оптимізації витрат дозволяє суттєво знизити простої, уникнути аварійних ситуацій та підвищити загальну ефективність підприємств зв'язку.

Інтелектуальні системи, що базуються на машинному навчанні, аналітиці даних і нейромережевих підходах, здатні своєчасно виявляти ризики відмов, пропонувати оптимальні сценарії заміни та адаптуватися до змін у технічному середовищі. Особлива увага в дослідженні приділена не лише високій точності прогнозів, а й їхній інтерпретованості (XAI), що є вирішальним фактором для довіри інженерного персоналу та фінансового обґрунтування рішень про оновлення.

У зв'язку з цим, дослідження, що спрямоване на розробку методу оптимізації процесу заміни обладнання з використанням ШІ, є актуальним як з наукової, так і з практичної точки зору. Такий підхід відповідає сучасним

тенденціям розвитку телекомунікаційної галузі та дозволяє забезпечити стабільне функціонування в умовах зростаючих вимог до якості та надійності зв'язку.

Мета роботи – удосконалення процесу заміни та оновлення обладнання на підприємствах зв'язку за допомогою розробленої методики з використанням штучного інтелекту.

Об'єкт дослідження – процес оновлення обладнання на підприємствах зв'язку.

Предмет дослідження – методика оновлення обладнання підприємств зв'язку з використанням прогностичної аналітики.

Для досягнення мети були вирішені наступні *завдання*:

1. Провести аналіз сучасного стану та існуючих методів оновлення обладнання на підприємствах зв'язку, виявити їхні недоліки.
2. Розробити математичну модель прогнозування технічного стану обладнання.
3. Сформуванати комплексну методику оновлення обладнання, що поєднує розроблену модель прогнозування та методи пояснювального штучного інтелекту.
4. Здійснити програмну реалізацію запропонованої методики.
5. Провести експериментальне дослідження ефективності розробленої системи, оцінити точність прогнозування та проаналізувати часові показники роботи модулів.

Наукова новизна роботи полягає в удосконаленні методики оновлення телекомунікаційного обладнання шляхом застосування оптимізованих ансамблевих моделей машинного навчання. Запропоновано комплексний підхід, який поєднує прогнозування залишкового ресурсу з класифікацією рівнів технічного ризику, що дозволяє адаптувати стратегію обслуговування до реального стану активів.

Практичне значення роботи полягає у створенні комплексного підходу, який поєднує прогнозування залишкового ресурсу з класифікацією рівнів

технічного ризику, що дозволяє адаптувати стратегію обслуговування до реального стану активів.

Результати досліджень апробовано шляхом публікації тез доповідей на наступних конференціях:

1. Пермякова О.О., Шевченко С.М. Застосування штучного інтелекту для оптимізації процесу заміни обладнання на підприємствах зв'язку: проблематика та рішення. V Всеукраїнська науково-практична конференція «Сучасні інтелектуальні інформаційні технології в науці та освіті». Збірник тез. 15.05.2025, ДУІКТ, м. Київ. К.: ДУІКТ, 2024, с.36

2. Пермякова О.О., Шевченко С.М. Безпека використання штучного інтелекту у процесах заміни обладнання на підприємствах зв'язку // Матеріали XII Всеукраїнської науково-практичної конференції молодих учених «Інформаційні технології – 2025», 15.05.2025, КСУ імені Бориса Грінченка, Київ, Україна, с.310-311

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Огляд та аналіз літературних джерел

У сучасному телекомунікаційному секторі, що характеризується стрімким переходом до технологій 5G, віртуалізації мережевих функцій (NFV) та Інтернету речей (IoT), проблема ефективного управління життєвим циклом інфраструктури набуває критичного значення. Традиційні підходи до заміни обладнання, що базуються на фіксованих термінах експлуатації або реагуванні на збої, стають економічно невиправданими через високу вартість простоїв та складність сучасних мереж. У зв'язку з цим, у світовій науковій спільноті спостерігається значне зростання інтересу до впровадження інтелектуальних систем підтримки прийняття рішень, заснованих на технологіях штучного інтелекту (ШІ) та машинного навчання. Аналіз останніх публікацій дозволяє виділити декілька ключових напрямків досліджень у цій галузі.

Фундаментальні зміни в підходах до обслуговування телеком-інфраструктури розглядаються у роботі Вассермана М. [1]. Автор аналізує вплив цифрової трансформації на процеси модернізації мереж, наголошуючи, що в умовах зростання трафіку ручне планування заміни обладнання стає неефективним. Вассерман акцентує увагу на необхідності переходу від моделі «ремонт після відмови» до предиктивної моделі, де ШІ виступає інструментом для прогнозування залишкового ресурсу обладнання. Це дозволяє не лише своєчасно виконувати заміну критичних вузлів, але й оптимізувати операційні витрати підприємства, уникаючи передчасного списання працездатних активів.

Важливий внесок у розробку методів автоматизованого контролю зроблено у вітчизняному дослідженні Завадського С. О. та Гуменюка Д. В. [2]. Автори зосередилися на оптимізації передачі даних у складних мультимедійних IoT-мережах. Хоча їх робота першочергово стосується маршрутизації трафіку, запропонований ними метод на базі ШІ може бути ефективно адаптований для

моніторингу стану телекомунікаційного обладнання. Аналізуючи метрики якості каналу (затримки, втрати пакетів), можна опосередковано визначати момент фізичного або морального зносу мережевих інтерфейсів, що є важливою складовою прогнозної аналітики.

Значна частина сучасних досліджень присвячена безпосередньо методології прогнозного обслуговування (Predictive Maintenance). У ґрунтовній оглядовій роботі Al-Turjman F. та співавторів [3] систематизовано сучасні підходи до впровадження прогнозного обслуговування у телекомунікаціях. Автори розглядають широкий спектр методів машинного навчання – від класичних регресійних моделей до складних ансамблевих алгоритмів. Ключовим висновком роботи є те, що ефективне планування заміни обладнання неможливе без урахування двох факторів: поточного технічного стану (Health Index) та інтенсивності навантаження на вузол.

Практичну реалізацію цих концепцій демонструють Haddara M. і Moen H. [4]. У своїй публікації вони описують розробку веб-застосунку, який використовує моделі машинного навчання для аналізу сенсорних даних у реальному часі. Дослідження доводить, що інтеграція аналітичних модулів у веб-інтерфейс дозволяє інженерам оперативно отримувати попередження про аварійні ситуації, що значно спрощує процес планування заміни та зменшує ризик людської помилки.

Специфіка телекомунікаційного обладнання полягає у генерації величезних обсягів телеметрії. У дослідженні Maharana S. та колег [5] зосереджено увагу саме на методах обробки потокових даних (stream processing) для виявлення аномалій. Використання алгоритмів ШІ дозволяє в режимі реального часу ідентифікувати навіть незначні відхилення у роботі обладнання (наприклад, поступове підвищення температури процесора або зростання кількості помилок на порту), які часто є передвісниками критичних збоїв. Це дає можливість прогнозувати необхідність заміни компонента ще до того, як він вплине на якість сервісу для абонента.

Стратегічний аспект заміни обладнання розглядається у статті Fernandez J. та співавтори [6]. Автори пропонують концепцію керування повним життєвим циклом активів (Asset Lifecycle Management) за допомогою інструментів ШІ. Акцент зроблено на задачі оптимізації: як вибрати ідеальний момент для заміни, щоб мінімізувати сукупну вартість володіння, але не допустити деградації послуг.

Цю тему розвиває робота Kumar D. та співавторів [7], де запропоновано модель оцінювання ризиків відмов. Використовуючи історичні дані про інциденти, автори навчають моделі ШІ формувати матрицю ризиків для різних типів обладнання. Це дозволяє підприємствам пріоритезувати заміну компонентів: спочатку замінюються вузли з найвищим ризиком відмови та найбільшим впливом на мережу, що забезпечує раціональний розподіл бюджету.

З розвитком обчислювальних потужностей все частіше застосовуються методи глибокого навчання. У дослідженні Farsi M. та співавторів [8] представлено підхід до побудови систем прогнозного обслуговування із використанням глибоких нейронних мереж. Автори демонструють, що нейромережі здатні знаходити приховані нелінійні залежності в даних телеметрії, які не помітні для традиційних статистичних методів. Це дозволяє суттєво підвищити точність прогнозування моменту відмови, що є критично важливим для оптимізації складних графіків заміни обладнання.

Інтеграція прогнозних моделей у бізнес-процеси висвітлюється у роботі Fauzi M. і співавторів [9]. Автори пропонують аналітичний фреймворк, який поєднує предиктивну аналітику з процесами логістики та закупівель. Запропонована ними модель дозволяє не просто передбачити збій, а й автоматично планувати постачання необхідних запчастин, зменшуючи час простою до мінімуму.

На завершення, варто відзначити вітчизняну статтю Ковальова Ю. та Ігнатенка О. [10], яка адаптує світовий досвід до українських реалій. Автори розглядають підходи до оцінювання технічного стану з урахуванням специфіки зносу обладнання в умовах вітчизняних мереж. Запропонована ними система підтримки прийняття рішень базується на комплексному аналізі даних та надає

рекомендації щодо доцільності ремонту або повної заміни устаткування, що є важливим кроком до автоматизації управлінських процесів на українських підприємствах зв'язку.

Проведений аналіз літератури свідчить, що, хоча окремі аспекти застосування ШІ в телекомунікаціях досліджені досить глибоко (зокрема, виявлення аномалій та нейромережеве прогнозування), існує потреба у створенні комплексної методики. Більшість розглянутих рішень фокусуються на технічному боці прогнозування, часто залишаючи поза увагою інтерпретацію результатів для персоналу та інтеграцію економічних факторів прийняття рішень.

1.2 Аналіз процесу технічного обслуговування та заміни обладнання на підприємствах зв'язку

Процес технічного обслуговування та заміни обладнання на підприємствах зв'язку є критично важливою складовою діяльності телекомунікаційних компаній, адже справність, надійність та актуальність обладнання безпосередньо впливають на якість наданих послуг, задоволеність клієнтів, а також фінансові показники підприємства.

На сучасному етапі розвитку телекомунікаційного сектора спостерігається стрімке зростання обсягів переданої інформації, зростання попиту на швидкісні послуги, а також постійне оновлення технологічних стандартів (наприклад, перехід від 4G до 5G). Це, у свою чергу, зумовлює необхідність регулярного оновлення мережевого та інфраструктурного обладнання, а також підвищення ефективності процесу технічного обслуговування.

Аналіз процесу обслуговування та заміни обладнання на підприємствах зв'язку дозволяє виокремити такі ключові етапи:

- моніторинг стану обладнання;
- виявлення та діагностика несправностей;
- ухвалення рішення щодо ремонту або заміни;
- безпосередньо проведення ремонтних чи замінних робіт;

- подальше тестування та введення в експлуатацію.

Традиційно ці процеси здійснювалися за участю людини на основі періодичних перевірок або за наявності скарг від споживачів. Проте такий підхід часто є реактивним і може призводити до простоїв мережі, незадоволення клієнтів та додаткових витрат.

Варто зазначити, що в умовах стрімкого зростання трафіку та переходу до конвергентних мереж, архітектура телекомунікаційних підприємств стає все більш гетерогенною. Це означає, що в одній мережі одночасно функціонує обладнання різних поколінь (Legacy-системи поряд з новітніми SDN-рішеннями) та різних виробників. Така ситуація значно ускладнює процес моніторингу та діагностики, оскільки кожен вендор надає власні, часто несумісні між собою інструменти управління.

Внаслідок цього, інженерний персонал змушений працювати в умовах інформаційного перевантаження, аналізуючи тисячі повідомлень про події щодня. Більшість цих повідомлень є інформаційними або дублюючими, і лише незначна частка свідчить про реальні інциденти. Людський фактор у такій ситуації стає «вузьким місцем»: оператор фізично не в змозі помітити кореляцію між незначним підвищенням температури на магістральному маршрутизаторі та зростанням рівня помилок на рівні агрегації, що відбувається через кілька годин. Саме тут виникає потреба в автоматизованих аналітичних системах, здатних виявляти приховані патерни деградації, які залишаються непомітними для людини при використанні традиційних порогових методів контролю.

У контексті цифрової трансформації підприємств зв'язку актуальним є впровадження інтелектуальних систем, зокрема технологій штучного інтелекту, які дають змогу переходити до прогностичного технічного обслуговування. Це означає, що завдяки аналізу даних отриманих з сенсорів обладнання, систем моніторингу та історичних даних про поломки, можна заздалегідь виявити ознаки потенційної несправності та здійснити заміну чи ремонт до настання критичної ситуації.

Таким чином, аналіз об'єкта дослідження дозволяє стверджувати, що його оптимізація за допомогою ШІ є перспективним напрямом, який здатен підвищити ефективність роботи підприємств зв'язку, скоротити витрати на обслуговування, зменшити кількість аварійних ситуацій та забезпечити вищу якість обслуговування користувачів.

1.3 Класифікація телекомунікаційного обладнання та фактори впливу на його надійність

Ефективність впровадження методики прогнозової аналітики неможлива без чіткого розуміння об'єктів управління. Телекомунікаційна мережа сучасного підприємства зв'язку являє собою складну гетерогенну систему, яка включає різноманітні класи апаратного забезпечення. Кожен клас має унікальну архітектуру, умови експлуатації та специфічні патерни деградації, що вимагає диференційованого підходу до прогнозування їхнього життєвого циклу.

Відповідно до загальноприйнятої трирівневої моделі побудови мереж (Three-Tier Network Architecture), розробленої компанією Cisco Systems, та рекомендацій Міжнародного союзу електрозв'язку (ITU-T), все активне обладнання можна розділити на три функціональні рівні, а саме: ядро мережі, рівень агрегації та рівень доступу.

1. Рівень ядра (Core Layer)

Це центральна частина мережі, яка відповідає за високошвидкісну комутацію пакетів та маршрутизацію магістрального трафіку. До обладнання цього рівня належать:

- магістральні маршрутизатори (Core Routers). Це високопродуктивні пристрої модульної архітектури, здатні обробляти трафік на швидкостях від 100 Гбіт/с до кількох Тбіт/с;
- комутатори ЦОД (Data Center Switches). Вони забезпечують зв'язок між серверами та системами зберігання даних;

– оптичні транспортні системи (DWDM). Це обладнання для ущільнення каналів зв'язку.

Обладнання ядра працює в умовах стабільного мікроклімату (гермо зони, прецизійні кондиціонери) з резервуванням живлення за схемою N+1 або 2N. Тому відмови тут рідко пов'язані з зовнішніми факторами. Критичними є внутрішні процеси, перегрів через високу щільність компонування та деградація високошвидкісних інтерфейсів.

2. Рівень розподілу/агрегації (Distribution/Aggregation Layer)

Цей рівень виступає проміжною ланкою, що збирає трафік з численних точок доступу та направляє його до ядра. Обладнання включає:

- L3-комутатори агрегації. Вони виконують маршрутизацію між VLAN, фільтрацію трафіку (ACL) та забезпечення якості обслуговування (QoS);
- сервери широкопasmового доступу (BRAS/BNG), які відповідають за авторизацію абонентів та шейпінг трафіку.

Це найбільш "стресовий" рівень з точки зору логічного навантаження. Процесори (CPU) та пам'ять (TCAM) цих пристроїв постійно завантажені обробкою таблиць маршрутизації та правил безпеки. Відмови часто мають змішаний характер, тобто апаратний знос накладається на програмні помилки.

3. Рівень доступу (Access Layer)

Це точка підключення кінцевих користувачів. Сюди входять:

- комутатори доступу (L2 Switches), які встановлюються у житлових будинках або офісах;
- базові станції мобільного зв'язку (eNodeB/gNodeB). Це радіобладнання для 4G/5G мереж;
- клієнтські термінали (ONU/CPE), які є пристроями на стороні абонента.

Обладнання часто розміщується в неконтрольованих умовах (вуличні шафи, горища, підвали), де на нього впливають екстремальні температури, вологість, пил та нестабільність електромережі. Це призводить до прискореного фізичного старіння компонентів.

Фактори впливу на надійність можна класифікувати на чотири групи (таблиця 1.1).

Таблиця 1.1

Класифікація факторів впливу на надійність телекомунікаційного обладнання

Група факторів	Фізичний механізм впливу	Характерні ознаки (Телеметрія)	Типові наслідки
Термічні	Температурна деградація. Згідно з законом Арреніуса, підвищення температури на 10°C прискорює хімічні реакції старіння вдвічі.	Висока температура CPU/Board, високі оберти вентиляторів (Fan RPM).	Висихання електrolітичних конденсаторів, відшарування підкладок чіпів (BGA), тепловий пробій транзисторів.
Електричні	Електрична ерозія та стрес. Мікрострибки напруги, перехідні процеси при комутації, електростатичні розряди.	Флуктуації напруги (Voltage), помилки на портах (CRC/FCS errors), падіння рівня оптичного сигналу (RX/TX Power).	Деградація лазерних діодів у SFP-модулях, вигорання портів, збої блоків живлення.
Механічні	Втома матеріалів та забруднення. Вібрації, накопичення пилу, окислення контактів.	Зниження ефективності охолодження при високих обертах вентиляторів, періодичні втрати лінку ("flapping").	Заклинювання вентиляторів, порушення контакту в роз'ємах RJ-45, перегрів через "пилу ковдру".
Логічні (Програмні)	Вичерпання ресурсів. Фрагментація пам'яті, помилки в мікрокоді (Firmware), перевантаження буферів.	Високе завантаження CPU (>90%), вичерпання RAM, відкидання пакетів (Drops), зростання затримок (Latency).	Зависання пристрою, спонтанні перезавантаження, відмова сервісів маршрутизації.

1.4 Аналіз існуючих рішень заміни обладнання на підприємствах зв'язку

Ефективність функціонування сучасного підприємства зв'язку безпосередньо залежить від обраної стратегії управління активами (Asset Management). Історично склалося так, що підходи до заміни та оновлення обладнання еволюціонували від найпростіших реактивних дій до складних планових процедур. На сьогоднішній день у практиці телекомунікаційних компаній співіснують декілька парадигм обслуговування, кожна з яких має свою сферу застосування, переваги та суттєві економічні обмеження. Детальний аналіз цих підходів дозволяє виявити системні недоліки, які неможливо усунути без застосування новітніх інформаційних технологій.

Варто зазначити, що в сучасних умовах гетерогенності мереж, коли в одній інфраструктурі співіснують технології різних поколінь (Legacy, 4G, 5G), обсяг даних телеметрії зростає експоненційно. Традиційні методи управління активами, розроблені для статичних телефонних мереж минулого століття, втрачають свою актуальність. Вони не враховують динамічну природу навантаження на мережеві вузли та вплив зовнішніх факторів середовища, що призводить до дисбалансу між надійністю надання послуг та вартістю їх забезпечення. Саме тому пошук нових, адаптивних механізмів оновлення стає не просто технічною задачею, а стратегічним пріоритетом для збереження конкурентоспроможності оператора.

На сьогодні в телекомунікаційній галузі превалюють три основні стратегії управління життєвим циклом обладнання, які еволюціонували від найпростіших до найбільш комплексних, проте все ще мають суттєві обмеження без застосування прогнозової аналітики: реактивний підхід, планово-попереджувальний підхід та обслуговування за фактичним станом.

1. Реактивне обслуговування (Breakdown Maintenance). Це найпростіший, але найбільш ризикований підхід, який домінував на підприємствах зв'язку у минулому. Обслуговування або заміна обладнання відбувається лише після того,

як сталася відмова. Перевага цього методу полягає у мінімальних початкових витратах на моніторинг. Проте недоліки є критичними: неконтрольовані простої мережі призводять до значних фінансових втрат, порушення угод про рівень обслуговування та втрати лояльності клієнтів. Окрім очевидних прямих збитків від простою, реактивна модель несе значні приховані витрати. По-перше, це штрафні санкції за порушення угод про рівень обслуговування (SLA) перед корпоративними клієнтами. По-друге, аварійна логістика запасних частин (експрес-доставка, понаднормова робота інженерів у нічний час) коштує в середньому на 30–50% дорожче за планову. Більше того, робота в режимі постійного "гасіння пожеж" призводить до швидкого професійного вигорання персоналу та зниження загальної якості експлуатації мережі. Такий підхід робить неможливим ефективне планування ресурсів та закупівлю запчастин.

2. Планово-попереджувальне обслуговування (Preventive Maintenance). Ця стратегія є більш поширеною, особливо для стандартизованих компонентів. Вона базується на моделі планової заміни з жорсткими інтервалами (наприклад, кожні 5 років або після X годин роботи), незалежно від фактичного технічного стану. Тобто задача зводиться до визначення оптимального терміну експлуатації, критерієм якого є мінімізація сумарних витрат на обслуговування протягом певного проміжку часу. Перевага полягає у зменшенні ймовірності раптових збоїв та спрощеному плануванні. Однак ключовий недолік – це надлишкові капітальні та операційні витрати, оскільки часто замінюється обладнання, яке має значний залишковий ресурс. Статистичний аналіз надійності складних технічних систем показує, що ймовірність відмови не завжди лінійно залежить від віку обладнання. Багато компонентів (наприклад, напівпровідникові плати маршрутизаторів) можуть стабільно працювати значно довше за гарантійний термін, якщо вони експлуатуються в ідеальних кліматичних умовах. Заміна такого обладнання виключно за календарем призводить до штучного завищення капітальних витрат. Фактично, підприємство сплачує за ресурс, який не був використаний, що знижує показник повернення інвестицій в інфраструктуру. Ця модель не враховує

реальну інтенсивність навантаження та вплив умов експлуатації (температура, вологість), що є критичним для підприємств зв'язку.

3. Обслуговування за фактичним станом (Condition-Based Maintenance, CBM). Цей підхід являє собою еволюцію попередніх методів. Він передбачає моніторинг телеметрії (навантаження процесора, температура, кількість помилок на портах) та діагностику несправностей. Рішення про ремонт або заміну приймається лише тоді, коли показники виходять за встановлені порогові значення. Це дозволяє здійснювати обслуговування лише за необхідності. Суттєвим недоліком CBM є проблема статичних порогових значень. Встановлення універсального порогу (наприклад, "завантаження CPU > 80%") для всієї мережі є неефективним, оскільки для ядра мережі такий показник може бути нормою в години пік, а для рівня доступу вже є ознакою критичного збою (наприклад, зависання процесу або DDoS-атаки). Це призводить до генерації великої кількості хибних сповіщень, через що оператори моніторингу часто ігнорують реальні інциденти, сприймаючи їх як "інформаційний шум". Метод CBM має суттєве обмеження: він покладається на людську інтерпретацію великих обсягів даних і, найголовніше, не дає прогностичної оцінки "коли" відбудеться відмова. Іншими словами, CBM сигналізує, що "стан поганий", але не може точно сказати, чи відбудеться критичний збій через 2 дні чи 2 тижні, що ускладнює своєчасне логістичне та фінансове планування.

На противагу традиційним підходам, останні наукові дослідження вказують на необхідність впровадження систем на базі ШІ, що поєднують аналіз стану за телеметрією з модулями прогностичної аналітики. У дослідженні Wang et al. [1] описано приклад впровадження інформаційної системи обліку з використанням нечіткої логіки для визначення часу заміни елементів оптичної мережі, яка дозволила зменшити витрати на 12% за рахунок відмови від надлишкових замін. Однак системи нечіткої логіки є обмеженими у здатності до самонавчання. Інші підходи базуються на побудові моделей прогнозування відмов (predictive failure models) із використанням алгоритмів класифікації, таких як Random Forest та XGBoost. Хоча ці моделі на практиці показали зменшення аварійних ситуацій,

вони часто не надають обґрунтування свого рішення, що створює проблему "чорної скриньки" і знижує довіру інженерного персоналу до автоматизованих рекомендацій. Таким чином, існує потреба у розробці комплексної методики, яка поєднує високу точність ШІ з прозорістю рішень. Порівняльний аналіз методів оновлення обладнання представлено у таблиці 1.2.

Таблиця 1.2

Порівняльний аналіз існуючих методів оновлення обладнання

Підхід	Метод	Прогно з відмови	Оптимізація	Управління запасами	Гнучкість (до змін у мережі)
Реактивне (Breakdown)	Ремонт після збою (Аварійний)	-	-	- (Аварійна закупівля)	-
Планово-попереджувальне (Preventive)	Заміна за календарем (Жорсткий інтервал)	-	- (Надлишкові заміни)	+	-
За фактичним станом	Моніторинг, порогові значення	-	+ (Краще, ніж планове)	- (Немає часу на замовлення)	+
Прогнозна аналітика	Розроблений метод	+	+	+	+(Адаптивна)

2 АНАЛІЗ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ РОЗРОБКИ МЕТОДИКИ

2.1 Штучний інтелект у прогностному обслуговуванні (Predictive Maintenance)

2.1.1 Основні переваги використання прогностної аналітики на підприємствах зв'язку

Впровадження технологій прогностної аналітики (Predictive Analytics) у виробничі процеси підприємств телекомунікаційної сфери є не просто технічним оновленням, а фундаментальним фактором цифрової трансформації бізнесу. Цей процес дозволяє здійснити перехід від архаїчних, реактивних моделей управління інфраструктурою до сучасної концепції прогностного обслуговування. Розглянемо детально ключові переваги цього підходу, які формують додану вартість для оператора зв'язку.

Традиційні підходи до експлуатації мереж мають суттєві економічні вади. Реактивний підхід (Run-to-Failure), що передбачає ремонт обладнання лише після його виходу з ладу, неминуче призводить до неконтрольованих витрат. Аварійні виїзди ремонтних бригад, особливо у нічний час або вихідні дні, оплачуються за підвищеними тарифами, а термінова логістика запасних частин вимагає значних коштів.

З іншого боку, планово-попереджувальний підхід (Preventive Maintenance), який базується на жорстких календарних графіках заміни, є економічно неефективним через проблему «недовикористаного ресурсу». Статистика свідчить, що значна частина компонентів (наприклад, вентилятори охолодження, акумуляторні батареї, оптичні трансивери), які замінюються за регламентом, могли б ефективно працювати ще тривалий час.

Прогнозна аналітика пропонує альтернативу – обслуговування «за станом» (Condition-Based Maintenance). Інтелектуальні алгоритми аналізують реальний

фізичний знос обладнання на основі даних телеметрії. Ключова перевага полягає в тому, що інженерні служби втручаються в роботу мережі виключно в той момент, коли модель ШІ сигналізує про високу ймовірність деградації або відмови. Це дозволяє:

- усунути витрати на обслуговування справного обладнання;
- зменшити кількість фізичних виїздів на об'єкти;
- оптимізувати використання робочого часу кваліфікованого персоналу.

Для телекомунікаційних компаній показник доступності послуг є критичним параметром, що регулюється жорсткими угодами про рівень обслуговування (Service Level Agreement). Будь-який, навіть короточасний, простій мережі призводить до прямих фінансових збитків (штрафні санкції, компенсації клієнтам) та, що не менш важливо, до втрати репутації та лояльності абонентів.

Моделі штучного інтелекту здатні ідентифікувати приховані патерни та аномалії в роботі обладнання задовго до настання критичного збою. Це дозволяє трансформувати потенційну аварію (яка вимагає негайної реакції в режимі «гасіння пожежі») у планове технічне завдання. Ремонтні роботи можуть бути заплановані у спеціальні технологічні «вікна», тобто періоди мінімального навантаження на мережу, зазвичай у нічний час. Такий підхід гарантує стабільність надання послуг та підвищує загальну якість сприйняття сервісу користувачем.

На стратегічному рівні впровадження прогностичної аналітики революціонує підхід до планування капітальних інвестицій. Технічні та фінансові департаменти отримують інструмент для точного розрахунку залишкового терміну служби кожного активу.

Замість масових закупівель нового обладнання для планової модернізації цілих сегментів мережі (що часто призводить до заморожування обігових коштів), компанія отримує можливість формувати гнучкий, адаптивний план оновлення. Це дозволяє:

- продовжити термін ефективної експлуатації існуючого парку обладнання;
- уникнути пікових навантажень на бюджет компанії;
- здійснювати закупівлі обладнання за оптимальними цінами, уникаючи термінових замовлень у постачальників.

Ще однією вагомою перевагою є перехід від складської моделі «Just-in-Case» (зберігання великих запасів «про всяк випадок») до моделі «Just-in-Time» (доставка компонентів під конкретну прогнозовану потребу).

Аналізуючи тренди зносу компонентів, система здатна заздалегідь сформулювати замовлення на необхідні запчастини. Це дозволяє суттєво розвантажити склади, зменшити витрати на зберігання та облік матеріальних цінностей, а також мінімізувати ризики морального старіння запасів, які роками лежать на складі і стають непотрібними ще до використання.

Таким чином, впровадження методики оновлення обладнання на основі прогнозної аналітики забезпечує комплексний позитивний ефект, що охоплює фінансові, технічні та організаційні аспекти діяльності підприємства зв'язку, створюючи надійний фундамент для його сталого розвитку в умовах висококонкурентного ринку.

2.1.2 Оцінка вимог до даних та очікуваної точності моделей

Успіх впровадження методики прогнозної аналітики фундаментально залежить від якості, гранулярності та повноти даних, що підтверджується принципом "Garbage In, Garbage Out". Для побудови ефективних моделей необхідна комплексна інтеграція різномірних джерел.

Насамперед, це історичні дані про відмови та обслуговування, які зазвичай зберігаються у системах Service Desk та базах даних управління конфігураціями. Ці дані містять критично важливу цільову змінну (факт відмови), хоча часто вони є неструктурованими (текстові описи проблеми), що вимагає застосування NLP для їх класифікації.

Другим компонентом є операційні дані моніторингу (телеметрія), що збираються в реальному часі. Для підприємств зв'язку це специфічні показники: лічильники помилок, параметри оптичних ліній, показники радіо середовища, дані SNMP (CPU load, memory usage) та syslog.

Третім компонентом є статичні дані про обладнання: модель, виробник, версія прошивки, дата введення в експлуатацію та умови середовища (наприклад, вулична шафа чи дата-центр). Критичною вимогою є забезпечення високої якості цих даних: очищення від аномалій, заповнення пропусків та, найголовніше, коректна синхронізація за часовими мітками для співставлення телеметрії з фактом відмови.

Щодо очікуваної точності, абсолютний 100% прогноз є недосяжною метою. Головна задача – досягти показників, що значно перевищують ефективність планових методів та мають прямий економічний ефект. Оцінка ускладнюється через сильний дисбаланс класів – стани "нормальної роботи" у тисячі разів переважають стани "відмови". В таких умовах метрика Ассурасу (загальна точність) є абсолютно неінформативною. Натомість, аналіз базується на матриці помилок (Confusion Matrix) та похідних метриках. Precision (точність) показує, який відсоток прогнозів "відмови" був вірним. Висока точність мінімізує хибні спрацьовування, тобто непотрібні виїзди інженерів та зайві витрати. Recall (повнота) показує, який відсоток реальних відмов модель змогла виявити. Висока повнота є критичною для мінімізації пропусків відмов, які є найбільш збитковими і ведуть до простою. Залежно від бізнес-стратегії (що є гіршим: даремно відправити інженера чи пропустити аварію), встановлюється баланс між цими метриками, який узагальнюється. На практиці, успішною вважається модель, що при прогнозуванні на заданому проміжку часу (наприклад, 7-14 днів) демонструє повноту та точність у діапазоні 85-95%, що є значним покращенням порівняно з традиційними підходами.

2.2 Огляд моделей машинного навчання для прогнозування

Усі існуючі підходи можна глобально розділити на три категорії: навчання з учителем (Supervised Learning), навчання без учителя (Unsupervised Learning) та навчання з підкріпленням (Reinforcement Learning).

У задачах прогнозного обслуговування найбільш поширеним є навчання з учителем. Цей підхід передбачає наявність історичного набору даних, де для кожного вектора ознак (вхідні параметри телеметрії) відома правильна відповідь (мітка класу стану: справний або несправний). Метою алгоритму є побудова функції відображення, яка з мінімальною похибкою апроксимує залежність між вхідними даними та цільовою змінною. До цього класу методів належать лінійна та логістична регресія, метод опорних векторів (SVM), дерева рішень та нейронні мережі.

Натомість, навчання без учителя застосовується у випадках, коли історичні дані про відмови відсутні або їх критично мало. У такому сценарії алгоритм намагається самостійно знайти структуру в даних, виявляючи кластери нормальної поведінки. Будь-яке суттєве відхилення від цих кластерів інтерпретується як аномалія. Типовими представниками цієї групи є метод головних компонент, автоенкодерів та метод ізоляційного лісу. Хоча ці методи є корисними для виявлення нових, раніше невідомих типів несправностей, вони часто страждають від високого рівня хибних спрацьовувань, оскільки будь-яка нестандартна поведінка мережі (наприклад, збільшення трафіку під час свят) може бути помилково сприйнята як технічний збій.

Для розробки методу оновлення обладнання було проаналізовано чотири основні класи моделей машинного навчання. Кожен з них вирішує специфічне завдання в рамках загального процесу прогнозного обслуговування, відрізняючись вимогами до даних, складністю реалізації та типом прогнозного результату.

1. Моделі регресії для прогнозування часу до відмови (Time-to-Failure)

Моделі регресії використовуються, коли цільова змінна y є неперервною величиною. У нашому завданні y – це "Залишковий термін служби" (RUL) або "Час до відмови", вимірний, наприклад, у днях або годинах роботи. Модель намагається знайти математичну функцію f , яка найкращим чином описує залежність між вхідними даними X (вектор ознак: температура, навантаження, вік, кількість помилок) та RUL.

Загальний вигляд моделі: $\hat{y} = f(X)$, де \hat{y} це прогнозований RUL.

Найпростішим прикладом є лінійна регресія, де залежність вважається лінійною. Модель шукає набір коефіцієнтів (ваг) β_i для кожної вхідної ознаки x_i :

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n, \quad (2.1)$$

де \hat{y} – прогнозований RUL;

x_i – значення i -ї ознаки (наприклад, x_i = 'середня температура');

β_i – вага, яку модель призначає i -й ознаці. Велика вага означає, що ця ознака сильно впливає на прогноз;

β_0 – вільний член (базове зміщення).

Модель "навчається", мінімізуючи функцію втрат. Для регресії найчастіше використовується Середньоквадратична помилка (Mean Squared Error, MSE), яка штрафувє модель за великі відхилення прогнозу \hat{y}_i від реального значення y_i :

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (2.2)$$

де N – кількість зразків у навчальній вибірці. Модель (наприклад, Random Forest Regressor) використовує більш складні нелінійні функції, але кінцева мета та сама – мінімізувати MSE.

Це найбільш стратегічно цінний тип прогнозу. Він дозволяє не просто реагувати, а оптимально планувати закупівлі, логістику запасних частин та час проведення робіт з оновлення, максимізуючи корисний час експлуатації обладнання.

Головна проблема – вимоги до даних. Для навчання потрібні "run-to-failure" датасети (дані, записані аж до самої відмови), які є рідкісними та дорогими.

Модель також погано працює, якщо відмови є раптовими і не мають чітких патернів деградації.

2. Моделі класифікації для визначення стану (Норма, Ремонт, Заміна)

Моделі класифікації використовуються, коли цільова змінна y є дискретною міткою (класом). У нашому випадку це класи: $y \in \{\text{'Норма'}, \text{'Ремонт'}, \text{'Заміна'}\}$. Модель не прогнозує коли, а прогнозує що. Вона вчиться знаходити в багатовимірному просторі ознак X розділові поверхні (decision boundaries), які найкращим чином відокремлюють зразки одного класу від іншого.

Замість прямого прогнозування класу, більшість сучасних класифікаторів (включно з деревами рішень та ансамблями) прогнозують ймовірність належності зразка X до кожного з k класів: $P = (y = j|X)$ для кожного класу j .

Наприклад, для зразка X_i (дані про комутатор) модель може видати:

$$P = (y = \text{'Норма'}|X_i) = 0.1$$

$$P = (y = \text{'Ремонт'}|X_i) = 0.7$$

$$P = (y = \text{'Заміна'}|X_i) = 0.2$$

Фінальним прогнозом \hat{y} обирається клас з найвищою ймовірністю (в даному випадку – 'Ремонт').

Для навчання моделей, що прогнозують ймовірності, використовується функція втрат Категоріальна крос-ентропія (Categorical Cross-Entropy). Вона вимірює, наскільки прогнозований розподіл ймовірностей \hat{y}_{ij} відрізняється від істинного y_{ij} (де істинна мітка – це 1 для правильного класу і 0 для всіх інших).

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k y_{ij} \log(\hat{y}_{ij}), \quad (2.3)$$

де N – загальна кількість зразків (одиниць обладнання) у навчальній вибірці;

k – кількість класів рекомендацій (наприклад, $k = 3$);

y_{ij} – істинна мітка (бінарний індикатор: 1, якщо зразок i належить до класу j , і 0 в іншому випадку);

\hat{y}_{ij} – прогнозована моделлю ймовірність того, що зразок i належить до класу j .

Цей підхід значно простіший в реалізації, оскільки зібрати та розмітити дані для класифікації легше, ніж повні "run-to-failure" дані. Моделі (особливо дерева рішень) часто є більш інтерпретованими – їхню логіку легко пояснити інженерному персоналу.

Результат є менш точним для стратегічного планування. Прогноз "Перед аварійний стан" сигналізує про проблему, але не дає точної відповіді, скільки часу залишилось, що ускладнює планування ресурсів.

3. Ансамблеві методи (Gradient Boosting) для підвищення точності прогнозу оновлення

Градiєнтний бустинг – це послідовна ансамблева техніка. На відміну від Random Forest, де дерева будуються паралельно та незалежно, бустинг будує дерева одне за одним, де кожне наступне дерево виправляє помилки попереднього. Процес виглядає так:

1) Будується перша проста модель F_0 (наприклад, одне дерево), яка робить початковий прогноз.

2) Розраховуються помилки (залишки) цієї моделі: $r_1 = \text{Істина} - \text{Прогноз}_0$.

3) Будується друге дерево h_1 , яке навчається прогнозувати не саму цільову змінну, а ці помилки r_1 .

4) Загальна модель оновлюється: $F_1 = F_0 + \alpha * h_1$ (де α – малий крок, "learning rate").

5) Розраховуються нові, менші помилки: $r_2 = \text{Істина} - \text{Прогноз}_1$.

6) Будується третє дерево h_2 , що прогнозує r_2 , і так далі.

Цей процес "навчання на помилках" є формою градієнтного спуску у просторі функцій. Кожне нове дерево h_m навчається на градієнті функції втрат (наприклад, MSE або Cross-Entropy) від попередньої моделі F_{m-1} . Загальна формула оновлення ансамблю на кроці m :

$$F_m(X) = F_{m-1}(X) + \alpha * h_m(X), \quad (2.4)$$

де $F_m(X)$ – фінальна модель ансамблю з m дерев;

$F_{m-1}(X)$ – модель з попереднього кроку;

$h_m(X)$ – нове дерево рішень, що "виправляє" помилки;

α – темп навчання (learning rate), що контролює внесок кожного нового дерева.

Зазвичай демонструють найвищу точність серед усіх методів машинного навчання на структурованих даних. Вони чудово обробляють змішані типи даних (числа, категорії), стійкі до викидів та здатні виявляти складні нелінійні залежності.

Є "чорною скринькою" (black box) – результат їхньої роботи складно інтерпретувати. Вони також вимагають ретельного налаштування гіперпараметрів і є більш обчислювально складними для навчання.

4. Моделі виявлення аномалій для моніторингу збоїв у реальному часі

Ці моделі базуються на навчанні без учителя (unsupervised). Їм не потрібні дані, розмічені як "відмова" ($y = 1$). Вони працюють на іншому принципі: аномалії є "малочисельними та відмінними". Це означає, що аномальний стан обладнання (наприклад, різкий стрибок температури) сильно відрізняється від мільйонів зразків "нормальної" роботи.

Алгоритм Isolation Forest (Ізолюючий Ліс):

1) Модель будує безліч випадкових дерев рішень (iTrees).

2) На відміну від звичайних дерев, iTrees не намагаються "вгадати" клас. Вони просто "ріжуть" простір даних випадковим чином: обирається випадкова ознака (наприклад, 'CPU load') і випадкове значення-поріг (наприклад, 'CPU load > 80%').

3) Ці поділи продовжуються, доки кожна точка даних (кожен зразок телеметрії) не опиниться в "ізоляції" (у своєму власному листі дерева).

Логіка:

– Нормальні точки (що лежать у щільних кластерах) вимагають багатьох випадкових поділів, щоб їх ізолювати. Вони матимуть велику глибину (довгий шлях) у дереві;

– Аномальні точки (що лежать далеко від інших) вимагають дуже мало поділів. Їх можна "відрізати" 1-2 поділами. Вони матимуть малу глибину (короткий шлях).

Модель розраховує середню довжину шляху $E(h(x))$ для кожної точки x по всіх деревах у лісі. На основі цього розраховується "оцінка аномальності" (s):

$$s(x, \varphi) = 2 \frac{E(h(x))}{c(\varphi)}, \quad (2.5)$$

де $E(h(x))$ – середня довжина шляху для зразка x ;

φ – кількість зразків, на яких будувалось дерево;

$c(\varphi)$ – коефіцієнт нормалізації (середня довжина шляху для дерева, побудованого на φ зразках).

Інтерпретація оцінки s :

- $s \approx 1$ (коли $E(h(x)) \approx 0$), тоді висока ймовірність аномалії.
- $s \approx 0.5$ (коли $E(h(x))$ середнє), тоді нормальна точка.
- $s \approx 0$ (коли $E(h(x))$ велике), тоді точно нормальна точка.

Головна перевага – не потрібні дані про відмови. Модель може працювати, маючи лише дані про нормальну експлуатацію. Це дозволяє виявляти нові, раніше невідомі типи збоїв.

Модель не дає жодної інтерпретації. Вона каже "щось не так", але не каже що саме не так і що з цим робити. Це часто призводить до високого рівня хибних спрацьовувань, де кожну аномалію інженеру доводиться перевіряти вручну.

5. Алгоритм випадкового лісу (Random Forest)

Суть методу полягає у побудові ансамблю незалежних дерев рішень, кожне з яких навчається на випадковій підвибірці вихідних даних.

Ключовою особливістю, що забезпечує високу ефективність Random Forest, є подвійна стохастичність. По-перше, для навчання кожного окремого дерева використовується бутстреп-вибірка (вибірка з поверненням), що дозволяє різним деревам фокусуватися на різних аспектах даних. По-друге, при розбитті кожного вузла дерева алгоритм обирає найкращу ознаку не з усіх доступних, а лише з випадкової підмножини ознак. Це дозволяє декоррелювати дерева між собою,

тобто, навіть якщо в наборі даних є одна дуже сильна ознака, вона не буде домінувати у всіх деревах, що робить загальну модель стійкою до перенавчання та викидів.

Фінальне рішення в задачі класифікації приймається шляхом мажоритарного голосування. Об'єкт відноситься до того класу, за який проголосувала більшість дерев. Такий підхід дозволяє значно зменшити дисперсію прогнозу порівняно з поодиноким деревом рішень, забезпечуючи високу стабільність результатів навіть на зашумлених даних телеметрії.

Порівняльний аналіз моделей представлено у таблиці 2.1.

Таблиця 2.1

Зведена таблиця порівняльного аналізу моделей

Критерій	Моделі регресії	Моделі класифікації (Стан)	Ансамблі (Random Forest)	Ансамблі (Gradient Boosting)	Моделі виявлення аномалій
Основна задача	Прогноз "КОЛИ?" " (кількість днів до відмови)	Прогноз "ЩО?" (клас стану: Норма, Ремонт)	Підвищення стабільності та зменшення дисперсії (Bagging)	Підвищення точності шляхом виправлення помилок (Boosting)	"ЦЕ НОРМАЛЬНО?" (пошук відхилень)
Вимоги до даних	Потрібні повні "run-to-failure" дані (дуже рідкісні)	Потрібні розмічені класи стану (легше зібрати)	Мінімальна підготовка, стійкість до викидів та незбалансованих даних	Чутливий до шумів, потребує ретельного налаштування	Потрібні лише дані про "нормальну" роботу

Продовження таблиці 2.1

Зведена таблиця порівняльного аналізу моделей

Критерій	Моделі Регресії	Моделі Класифікації (Стан)	Ансамблі (Random Forest)	Ансамблі (Gradient Boosting)	Моделі Виявлення Аномалій
Переваги	Найвища цінність для стратегічного планування	Легкість інтерпретації, простота реалізації	Стійкість до перенавчання, висока швидкість навчання (паралельність)	Найвища точність прогнозу на табличних даних	Виявляє нові типи збоїв, не потребує даних про відмови
Недоліки	Висока складність, високі вимоги до якості даних	Менша цінність для планування, результат дискретний	Займає багато пам'яті, повільніший час прогнозу (inference)	"Чорна скринька", схильність до перенавчання, повільне навчання	Високий рівень хибних тривог, відсутність інтерпретації

2.3 Інтерпретація моделей та підтримка прийняття рішень

2.3.1 Методи пояснювального ШІ (XAI) для обґрунтування рекомендацій

Сучасні високоточні моделі, такі як ансамблі Gradient Boosting (XGBoost) або нейронні мережі, за своєю природою є "чорними скриньками" (black boxes). Вони надають точний прогноз, але не пояснюють, чому було прийняте саме таке рішення. Для обґрунтування дорогих рішень, як-от оновлення або заміна

обладнання, критично важливо розуміти логіку моделі. Для цього застосовуються методи пояснювального ШІ (Explainable AI, XAI).

XAI-методи поділяються на дві категорії:

1. Глобальна інтерпретація (Global Interpretability)

Ці методи пояснюють поведінку моделі в цілому. Найпоширенішим є аналіз важливості ознак (Feature Importance). Цей підхід розраховує, який відносний внесок кожна вхідна ознака (наприклад, вік обладнання, середня температура, кількість помилок) робить у загальну точність моделі. Це дозволяє на стратегічному рівні зрозуміти, які параметри є ключовими індикаторами зносу обладнання на підприємстві.

2. Локальна інтерпретація (Local Interpretability)

Це найбільш цінний підхід для обґрунтування конкретної рекомендації (наприклад, "Замінити комутатор №123"). Провідними методами тут є LIME (Local Interpretable Model-agnostic Explanations) та SHAP (SHapley Additive exPlanations).

SHAP є найбільш потужним з них. Базуючись на принципах теорії ігор (значеннях Шеплі), SHAP розраховує точний внесок кожної ознаки у фінальний прогноз для кожного окремого прогнозу.

Приклад: Модель прогнозує для комутатора №123 "ймовірність заміни = 90%".

Інженер отримує не лише цей прогноз, а й пояснення від SHAP:

- базова ймовірність (середня по всіх пристроях): 30%
- SHAP внесок 'вік = 10 років' (замість 4): +25%
- SHAP внесок 'кількість CRC-помилки = 1500/добу' (замість 20): +40%
- SHAP внесок 'температура = 70°C' (замість 55°C): +10%
- SHAP внесок 'версія ПЗ = стара': +5%
- Інші фактори...
- Разом: 90%

Таким чином, ХАІ перетворює непрозорий прогноз на обґрунтовану рекомендацію, надаючи інженеру чіткі докази, на основі яких дані модель дійшла свого висновку.

2.3.2 Підходи до візуалізації прогнозів для інженерного персоналу

Для того, щоб прогнози моделі ШІ приносили реальну користь, вони мають бути представлені кінцевим користувачам (інженерам, керівникам відділів експлуатації) у зрозумілому, інтерактивному та дієвому форматі. Сирі дані чи таблиці з ймовірностями для цього не підходять. Найкращим рішенням є розробка спеціалізованих дашбордів (інформаційних панелей).

Такий дашборд є системою підтримки прийняття рішень і повинен містити декілька ключових компонентів:

1. Високорівневий огляд

Це екран для менеджменту. Він зазвичай містить карту або таблицю всього парку обладнання, де кожна одиниця маркована кольором відповідно до прогнозованого ризику (наприклад, зелений – 'Норма', жовтий – 'Потрібен ремонт в термін до 30 днів', червоний – 'Критичний стан').

2. Пріоритетний список завдань

Це основний інструмент для інженерного персоналу. Він являє собою відсортований список обладнання, що вимагає негайної уваги. Список сортується за рівнем ризику (найвища ймовірність відмови) або за прогнозованим (від найменшого).

3. Деталізована картка активу

При натисканні на конкретну одиницю обладнання зі списку, користувач потрапляє на її "паспорт". Цей екран є критично важливим і має містити:

- чіткий прогноз. "Рекомендовано: Заміна" або "Прогнозований: 12 днів";
- обґрунтування ХАІ. Простий список ключових факторів, що вплинули на прогноз;

- історичні тренди. Графіки, що показують динаміку ключових показників (тих самих, що згадуються в ХАІ) за останній місяць. Це дозволяє інженеру візуально підтвердити, що, наприклад, кількість помилок дійсно стрімко зростає;

- довідкова інформація. Модель, серійний номер, локація, дата встановлення, історія минулих ремонтів.

Мета такого дашборду – перетворити складний прогноз ШІ на просте і зрозуміле завдання для інженера, відповідаючи на три питання: ЩО трапиться? ЧОМУ модель так вважає? КОЛИ потрібно втрутитись?

2.4 Проблеми обробки незбалансованих даних

Однією з ключових проблем при розробці систем прогнозування аналітики для підприємств зв'язку є специфічна природа вхідних даних, що характеризується суттєвим дисбалансом класів (Class Imbalance). У контексті експлуатації телекомунікаційного обладнання це означає, що кількість записів про нормальну роботу пристроїв (клас "Норма") у тисячі разів перевищує кількість записів про аварійні ситуації або відмови (клас "Заміна" або "Ремонт").

У типовому наборі даних телеметрії співвідношення класів може сягати 1000:1 або навіть 10000:1 на користь "здорового" обладнання. Якщо застосовувати стандартні алгоритми машинного навчання до таких даних без попередньої підготовки, виникає ефект, відомий як "Парадокс точності" (Accuracy Paradox).

Модель, прагнучи мінімізувати загальну помилку, може просто класифікувати всі зразки як "Норма". У такому випадку формальна точність (Accuracy) буде надзвичайно високою (наприклад, 99.9%), але практична цінність такої моделі дорівнюватиме нулю, оскільки вона не здатна виявити жодної реальної загрози (Recall = 0%).

Для підприємства зв'язку вартість помилок першого та другого роду є нерівнозначною. При False Positive (Хибна тривога) система помилково

рекомендує заміну справного обладнання. Це призводить до зайвих витрат часу інженера на перевірку, але не впливає на якість послуг. В той час при False Negative (Пропуск події) система класифікує аварійний пристрій як справний. Це призводить до раптової відмови, простою мережі та порушення SLA, що має значно тяжчі фінансові та репутаційні наслідки.

Тому стандартні функції втрат, які симетрично штрафують за обидва типи помилок, є неприйнятними. Необхідне застосування спеціалізованих методів балансування даних та модифікація алгоритмів навчання.

Основним підходом до вирішення проблеми дисбалансу є штучна зміна співвідношення класів у навчальній вибірці (ресемплінг). Існує два основні стратегічні напрямки:

1. Undersampling (Видалення прикладів мажоритарного класу). Це випадкове або алгоритмічне видалення частини записів про "нормальну" роботу. Цей метод дозволяє вирівняти баланс і прискорити навчання (через зменшення обсягу даних), але несе ризик втрати важливої інформації про рідкісні патерни нормальної поведінки.

2. Oversampling (Дублювання прикладів міноритарного класу). Це штучне збільшення кількості записів про аварії. Найпростіший варіант, просте дублювання рядків, часто призводить до перенавчання моделі, оскільки вона просто запам'ятовує конкретні приклади, а не вчиться узагальнювати ознаки аварійності.

Більш ефективним рішенням є генерація синтетичних даних, які статистично схожі на реальні аварії, але не є їх точними копіями. Метод SMOTE (Synthetic Minority Over-sampling Technique) є золотим стандартом у боротьбі з дисбалансом. Алгоритм працює у просторі ознак, а не з самими об'єктами. Для кожного прикладу з міноритарного класу (аварії) він знаходить найближчих сусідів з того ж класу. Потім на відрізку, що з'єднує вихідний приклад та його сусіда, випадковим чином обирається точка, яка стає новим синтетичним прикладом.

Математично генерація нового зразка x_{new} описується формулою:

$$x_{new} = x_i + \lambda * (x_{zi} - x_i), \quad (2.6)$$

де x_i – існуючий зразок міноритарного класу;

x_{zi} – один із його найближчих сусідів;

λ – випадкове число в діапазоні $[0, 1]$.

Метод ADASYN (Adaptive Synthetic Sampling) є удосконаленням SMOTE. Він адаптивно генерує більше синтетичних даних для тих прикладів міноритарного класу, які важче вивчити (тобто тих, що знаходяться близько до кордону з мажоритарним класом). Це дозволяє змістити межу прийняття рішень у бік проблемних зон, покращуючи здатність моделі розпізнавати неочевидні випадки деградації обладнання.

Окрім маніпуляцій з даними, ефективним підходом є модифікація самого процесу навчання. Навчання, чутливе до вартості (Cost-Sensitive Learning). Суть методу полягає у введенні вагових коефіцієнтів у функцію втрат моделі. Помилкам класифікації міноритарного класу присвоюється значно більша вага, ніж помилкам мажоритарного класу. Для ансамблевих моделей, таких як Random Forest або XGBoost, це реалізується через параметри зважування. Наприклад, якщо співвідношення класів становить 1:100, то вага помилки для класу "Заміна" встановлюється у 100 разів вищою. Це змушує алгоритм приділяти пріоритетну увагу правильній класифікації саме аварійних ситуацій, навіть якщо це призведе до деякого зниження загальної точності за рахунок збільшення кількості хибних тривог.

3 РОЗРОБКА МЕТОДУ ОНОВЛЕННЯ ОБЛАДНАННЯ НА ПІДПРИЄМСТВІ ЗВ'ЯЗКУ З ВИКОРИСТАННЯМ ПРОГНОЗНОЇ АНАЛІТИКИ

3.1 Математична модель прогнозування оновлення обладнання

Для розробки методики оновлення обладнання підприємств зв'язку використовується підхід на основі штучного інтелекту, зокрема ансамблевих моделей машинного навчання. Модель повинна приймати вхідні дані про стан обладнання та на їх основі прогнозувати необхідність його оновлення (ремонту або заміни).

1. Вхідні дані

Вхідні дані X являють собою вектор параметрів, що характеризують кожен одиницю обладнання на підприємстві зв'язку:

$X = \{x_1, x_2, \dots, x_n\}$, де x_i – параметри обладнання (кількісні та категоріальні).

До основних вхідних параметрів (x_i) належать:

- технічні характеристики (модель обладнання, виробник, рік введення в експлуатацію, нормативний термін служби);
- експлуатаційні показники (середнє навантаження, обсяг трафіку, час безвідмовної роботи);
- дані моніторингу (кількість збоїв або помилок за період, дані з датчиків, якщо є);
- історія обслуговування (кількість проведених ремонтів, дата останнього технічного обслуговування, загальна вартість обслуговування);
- економічні показники (вартість нового аналогічного обладнання, вартість ремонту).

2. Вихідні дані (Цільова змінна)

Цільова змінна R – це прогнозована рекомендація щодо конкретної одиниці обладнання. Вона визначається як задача класифікації.

$R \in \{R_1, R_2, \dots, R_k\}$, де R – прогнозована дія (рекомендація).

Наприклад, $k = 3$ і цільові класи можуть бути:

- R_1 = 'Експлуатація' (обладнання в нормі, втручання не потрібне);
- R_2 = 'Ремонт' (необхідне планове/позапланове обслуговування);
- R_3 = 'Заміна' (обладнання рекомендовано до повного оновлення).

3. Модель прогнозування

Для вирішення задачі класифікації стану обладнання використовується ансамблева модель, що складається з m дерев рішень (наприклад, модель градієнтного бустингу XGBoost або Random Forest). Ці моделі ефективно працюють зі змішаними (кількісними та категоріальними) даними, що характерно для цієї задачі.

Кожне j -те дерево в ансамблі $h_j(X)$ ($j = 1, 2, \dots, m$) видає ймовірнісний розподіл належності вхідного вектора X до одного з k класів. Результат ансамблю (фінальна рекомендація \hat{R}) формується шляхом вибору класу, який має максимальну сумарну (або усереднену) ймовірність по всіх деревах. Це зображено у формулі 3.1:

$$\hat{R} = \operatorname{argmax}_{R \in \{R_1, \dots, R_k\}} \left(\sum_{j=1}^m P_j(R|X) \right), \quad (3.1)$$

де \hat{R} – прогнозована рекомендація (клас);

$P_j(R|X)$ – ймовірність рекомендації R (наприклад, 'Заміна') за даними X , прогнозована j -м деревом.

4. Функція втрат

Для навчання моделі класифікації (оцінки розбіжностей між істинними мітками R та прогнозованими ймовірностями \hat{R}) використовується категоріальна крос-ентропія (Categorical Cross-Entropy). Вона показує, наскільки добре прогнозований розподіл ймовірностей відповідає істинному класу. Для врахування дисбалансу вона трансформується у зважену категоріальну крос-

ентропію (Weighted Categorical Cross-Entropy). Суть модифікації полягає у введенні вагового коефіцієнта α_j для кожного класу j .

Формула функції втрат L має вигляд (формула 3.2):

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k \alpha_j y_{ij} \ln(\hat{y}_{ij}), \quad (3.2)$$

де N – загальна кількість зразків (одиниць обладнання) у навчальній вибірці;

k – кількість класів рекомендацій (наприклад, $k = 3$);

α_j – вага класу j . Для рідкісних класів вага встановлюється обернено пропорційною їх частоті у вибірці;

y_{ij} – істинна мітка (бінарний індикатор: 1, якщо зразок i належить до класу j , і 0 в іншому випадку);

\hat{y}_{ij} – прогнозована моделлю ймовірність того, що зразок i належить до класу j .

Завдяки множнику α_j , навіть одна помилка при розпізнаванні аварійного стану створює велике значення втрат L .

3.2 Нормалізація даних

Етап попередньої обробки даних (Data Preprocessing) є критично важливим у будь-якій роботі з машинним навчанням, оскільки якість вхідної інформації безпосередньо визначає гранично досяжну точність моделі. У сфері телекомунікацій дані, що надходять із систем моніторингу, часто характеризуються неповнотою, зашумленістю та наявністю артефактів.

Типовою проблемою є пропуски в даних, які можуть виникати через тимчасову втрату зв'язку з пристроєм, перезавантаження контролера моніторингу або програмні помилки при агрегації логів. Просто видаляти такі записи не завжди доцільно, оскільки це може призвести до втрати цінної інформації,

особливо якщо пропуски не є випадковими (наприклад, пристрій перестає надсилати метрики безпосередньо перед відмовою).

Іншою важливою проблемою є аномальні викиди – значення, що суттєво відрізняються від загального розподілу. У телеметрії викиди можуть бути як ознакою реальної проблеми (стрибок напруги), так і результатом помилки вимірювання сенсора. Для коректної роботи алгоритмів необхідно застосовувати методи фільтрації, такі як правило трьох сигм або міжквартильний розмах (IQR), щоб мінімізувати вплив вимірювального шуму на процес навчання моделі.

Перед етапом моделювання вхідні дані потребують спеціальної підготовки, оскільки параметри телекомунікаційного обладнання мають різну природу та діапазони значень. Наприклад, температура процесора вимірюється в градусах Цельсія (діапазон 40–90°C), рівень помилок – у проміле або відсотках (0–5%), а час з моменту останнього обслуговування – у днях (0–365). Якщо подати такі "сирі" дані безпосередньо в алгоритми машинного навчання, ознаки з більшими числовими значеннями (наприклад, час у днях) будуть домінувати над ознаками з меншими значеннями (наприклад, рівень помилок), що спотворить результати навчання та знизить точність прогнозування.

Для вирішення цієї проблеми у методиці застосовується процедура нормалізації даних методом стандартизації. Цей метод трансформує значення кожної числової ознаки таким чином, щоб вони мали середнє значення (μ), рівне 0, та стандартне відхилення (σ), рівне 1.

Формула перетворення для кожного значення x_i виглядає наступним чином:

$$x'_i = \frac{x_i - \mu}{\sigma}, \quad (3.3)$$

де x'_i – нормалізоване значення параметра, яке подається на вхід моделі;

x_i – вихідне (сире) значення параметра;

μ – середнє арифметичне значення ознаки по всій навчальній вибірці;

σ – стандартне відхилення ознаки.

Використання стандартизації забезпечує рівний ваговий внесок усіх параметрів у роботу ансамблевої моделі, покращує збіжність алгоритмів оптимізації (наприклад, градієнтного спуску в XGBoost) та підвищує загальну стабільність системи прогнозування.

3.3 Алгоритм для аналізу стану обладнання

Для реалізації методики прогнозування, описаної в попередньому пункті (3.1), було обрано методи машинного навчання, які забезпечують високу точність у задачах класифікації та прогнозування зносу. Основними підходами стали дерева рішень та ансамблеві моделі. Дерева рішень дозволяють візуально інтерпретувати ключові фактори (наприклад, "кількість збоїв > N" або "термін експлуатації > Y років"), що є важливим для обґрунтування управлінських рішень щодо оновлення.

Особливу роль у запропонованій методиці відіграють ансамблеві моделі:

- Random Forest (Випадковий ліс) поєднує безліч дерев рішень для підвищення точності та, що важливо, стійкості моделі до "шумів" у даних (наприклад, аномальних показників навантаження чи поодиноких збоїв);
- Gradient Boosting (XGBoost або LightGBM) використовується для послідовного підсилення моделі. Кожна наступна модель (дерево) виправляє помилки попередніх, що дозволяє досягти найвищої точності прогнозування класу обладнання ('Експлуатація', 'Ремонт' чи 'Заміна');
- Stacking (Стекінг), як метод комбінування, може бути використаний для інтеграції результатів різних алгоритмів (наприклад, логістичної регресії та бустингу) в єдиний, більш надійний прогноз.

Ці підходи забезпечують необхідну точність та ефективність при роботі з багатовимірними даними про стан телекомунікаційного обладнання.

Розробка методики враховувала обмеження щодо обробки даних у реальному часі та неоднорідність інформації (поєднання технічних логів, фінансових показників та історії обслуговування). Для їх подолання були

застосовані методи оптимізації алгоритмів через налаштування гіперпараметрів та вибір ансамблевих методів, які ефективно обробляють великі обсяги різнотипних даних і забезпечують високий рівень точності прогнозу.

Загальний алгоритм реалізації методики прогнозування зображено на рисунку 3.2.

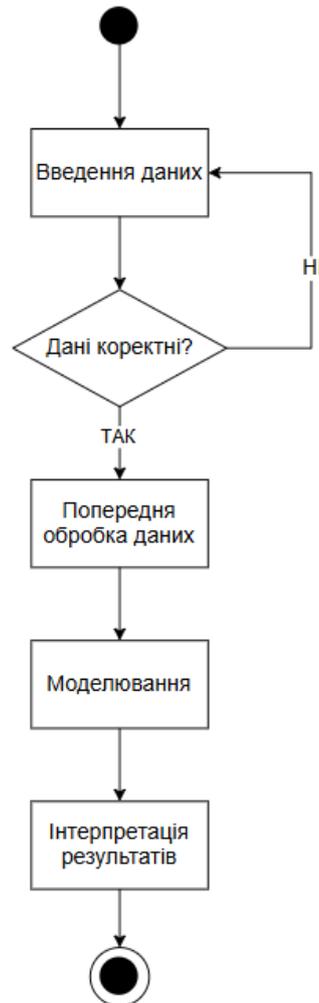


Рис. 3.1 Алгоритм реалізації методики прогнозування

На першому етапі, «Введення даних», система ініціює збір інформації про конкретну одиницю телекомунікаційного обладнання. Джерелом цих даних можуть виступати системи мережевого моніторингу (NMS) або бази даних інвентаризації. Вхідний вектор параметрів включає динамічні показники телеметрії, такі як середнє навантаження на процесор (CPU load), температура компонентів, рівень помилок на інтерфейсах, а також статичні характеристики,

зокрема вік обладнання та час, що минув з моменту останнього технічного обслуговування.

Одразу після отримання даних відбувається етап «Перевірка коректності даних». Це критично важливий крок валідації, на якому система аналізує вхідний вектор на предмет повноти та відповідності фізичним обмеженням. Якщо виявлено пропущені значення або аномалії, що свідчать про збій моніторингу (наприклад, від'ємна температура або некоректний формат дати), процес переривається, і система повертається до етапу введення для повторного запиту. Це гарантує, що на вхід моделі машинного навчання потрапляють лише якісні та верифіковані дані.

Після успішної валідації розпочинається етап «Попередня обробка даних». Оскільки алгоритми машинного навчання чутливі до масштабу та типу даних, цей модуль виконує дві основні трансформації. По-перше, числові показники (температура, навантаження) нормалізуються за допомогою методу стандартизації (StandardScaler), що приводить їх до єдиного діапазону значень. По-друге, категоріальні змінні, такі як тип або модель пристрою, кодуються у бінарний векторний формат за допомогою методу One-Hot Encoding.

Центральним етапом алгоритму є «Моделювання». На цьому кроці підготовлений вектор даних подається на вхід попередньо навченого ансамблевого класифікатора (наприклад, Random Forest). Модель виконує математичний аналіз сукупності факторів і прогнозує поточний клас технічного ризику обладнання: «Нормальна експлуатація», «Плановий ремонт» або «Термінова заміна». Окрім самого класу, модель також розраховує рівень впевненості (ймовірність) цього прогнозу, що слугує додатковим індикатором надійності.

Завершальним етапом є «Інтерпретація результатів». На відміну від традиційних систем, які видають лише сухий прогноз, розроблена методика використовує модуль пояснювального штучного інтелекту (XAI). Система аналізує вхідні фактори, що найбільше вплинули на рішення моделі (наприклад, поєднання критичного віку та перегріву), і автоматично формує текстове

обґрунтування. У результаті інженер отримує комплексну рекомендацію, яка містить не лише вказівку до дії, а й прозоре пояснення причин, що дозволяє приймати обґрунтовані рішення щодо оновлення інфраструктури.

3.4 Застосування математичної моделі

3.4.1 Засоби розробки та архітектура ПЗ

Реалізація математичної моделі, описаної в попередніх підрозділах, розпочалася з обґрунтування та вибору інструментарію розробки. В якості основного середовища було обрано високорівневу мову програмування Python, яка на сьогодні є стандартом де-факто у сфері науки про дані та машинного навчання. Вибір зумовлений наявністю потужної екосистеми спеціалізованих бібліотек, які дозволяють ефективно реалізувати весь процес обробки даних – від завантаження логів в їх початковому стані до візуалізації прогнозів.

Ключовим компонентом для маніпуляцій з даними стала бібліотека Pandas, яка забезпечує високу продуктивність при роботі з часовими рядами телеметрії та дозволяє виконувати складні операції агрегації та фільтрації.

Для реалізації алгоритмів машинного навчання використано бібліотеку Scikit-learn, яка містить необхідні інструменти для попередньої обробки даних, масштабування ознак та побудови базових моделей.

Для підвищення точності прогнозування додатково інтегровано бібліотеки градієнтного бустингу XGBoost та LightGBM, які демонструють найкращі результати на табличних даних.

Серіалізація навчених моделей здійснюється за допомогою бібліотеки Joblib, що дозволяє зберігати стан моделі та використовувати її в промисловому середовищі без необхідності витратного за часом повторного навчання.

Архітектура програмного комплексу спроектована за об'єктно-орієнтованим принципом, що забезпечує модульність коду та легкість його подальшого масштабування. Система складається з кількох взаємопов'язаних класів, кожен з яких відповідає за окремий етап обробки інформації. Клас TelemetryLoader

відповідає за шар доступу до даних, зчитуючи інформацію з CSV-файлів або безпосередньо підключаючись до баз даних систем моніторингу. Отримані дані передаються до класу `DataPreprocessor`, який виконує критично важливі функції очищення від шумів, обробки пропущених значень та генерації нових ознак. Саме тут відбувається трансформація початкових показників датчиків у значущі для моделі характеристики, такі як динаміка зміни температури або частота виникнення помилок за певний період.

Ядром системи виступає клас `ModelEngine`, який інкапсулює логіку машинного навчання. Він містить методи для тренування обраного алгоритму (наприклад, `Random Forest`), виконання крос-валідації та оптимізації гіперпараметрів. Завершальним елементом архітектури є клас `DecisionSupport`, який трансформує математичні ймовірності, отримані від моделі, у зрозумілі бізнес-рекомендації для персоналу. Для взаємодії з кінцевим користувачем – інженером з експлуатації – розроблено графічний інтерфейс на базі бібліотеки `Tkinter` або `CustomTkinter`, який візуалізує результати аналізу у зручному для сприйняття вигляді.

3.4.2 Програмна реалізація алгоритму

На початковому етапі програмної реалізації було створено модуль підготовки даних, який вирішує проблему гетерогенності вхідної інформації. Оскільки дані про телекомунікаційне обладнання містять не лише числові показники телеметрії, але й категоріальні змінні, такі як тип пристрою (маршрутизатор, комутатор, шлюз) або назва моделі, було застосовано метод `One-Hot Encoding`. Цей підхід дозволив перетворити текстові назви у бінарні вектори, які коректно сприймаються алгоритмами машинного навчання, не створюючи при цьому хибної ієрархії між категоріями. Також на цьому етапі проводиться нормалізація числових даних, що є критичним для коректної роботи багатьох алгоритмів.

Наступним кроком стало навчання моделі, для чого підготовлений набір даних було розділено на тренувальну та тестову вибірки у співвідношенні 80/20 за

допомогою функції `train_test_split`. Це стандартна практика, яка дозволяє об'єктивно оцінити здатність моделі узагальнювати виявлені закономірності на нових даних, які вона не бачила під час навчання. Під час тренування особливу увагу було приділено боротьбі з перенавчанням та дисбалансом класів, що часто зустрічається в задачах прогнозування відмов, де кількість справного обладнання значно перевищує кількість несправного.

Ключовим етапом підвищення ефективності системи стала оптимізація гіперпараметрів обраної моделі `Random Forest`. Було експериментально встановлено, що критичний вплив на точність прогнозу мають такі параметри, як кількість дерев у ансамблі (`n_estimators`) та їх максимальна глибина (`max_depth`). Замість ручного підбору було реалізовано автоматизований підхід з використанням методу `GridSearchCV`, який перебирає задані комбінації параметрів та оцінює їх ефективність за допомогою крос-валідації. Це дозволило знайти оптимальну конфігурацію моделі та підвищити точність прогнозування на 5-7% порівняно з базовими налаштуваннями за замовчуванням.

Результатом роботи алгоритму є не просто ймовірнісна оцінка, а чітка інструкція для інженера. Система аналізує ймовірність відмови і, якщо вона перевищує встановлений поріг, формує картку інциденту. Наприклад, для об'єкта з високим ризиком відмови виводиться його поточний статус, розрахована ймовірність збою у відсотках та, що найважливіше, критичний фактор, який вплинув на такий прогноз, наприклад, стабільний перегрів процесора протягом останніх годин. На основі цього генерується конкретна рекомендація, наприклад, запланувати заміну системи охолодження у визначений термін. Такий підхід дозволяє технічному спеціалісту не витратити час на ручний аналіз тисяч рядків логів, а одразу отримувати готове рішення для запобігання аварії.

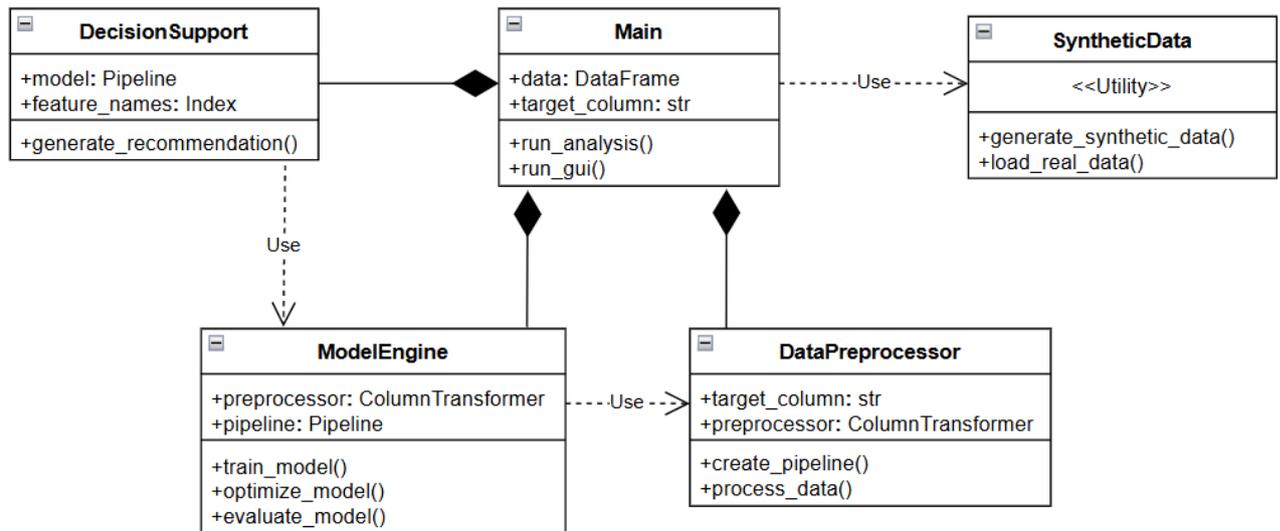


Рис. 3.2 Діаграма класів розробленої системи

Архітектура розробленої системи базується на об'єктно-орієнтованому підході, що забезпечує модульність, легкість підтримки та можливість масштабування. Система складається з п'яти основних класів:

1. Клас «Main» (Головний керуючий модуль)

Цей клас є точкою входу в програму та виступає в ролі "керуючого". Він відповідає за ініціалізацію системи, зберігання глобального стану даних та координацію взаємодії між модулями обробки, моделювання та візуалізації.

Атрибутами є `data` та `target_column`. Перший зберігає завантажений набір даних (історичні логи або поточну телеметрію) у форматі структурованої таблиці Pandas, другий це назва стовпця з цільовою змінною (класом стану: R1, R2, R3), яку система має прогнозувати.

Методами класу є `run_analysis()` та `run_gui()`. Перший запускає повний цикл аналітичної обробки (Back-end логіка). Метод послідовно викликає завантаження даних, їх попередню обробку, навчання моделі, оптимізацію гіперпараметрів та розрахунок метрик точності. Другий ініціалізує графічний інтерфейс користувача (Front-end на базі бібліотеки Tkinter), де відображаються результати роботи, метрики та картки з рекомендаціями для інженера.

2. Клас DataPreprocessor (Попередня обробка даних)

Відповідає за підготовку "сирих" технічних даних до формату, придатного для обробки алгоритмами машинного навчання.

Атрибутами є `target_column` та `preprocessor`. Перший це назва цільової змінної, яку необхідно виключити з масиву ознак під час трансформації. Другий це об'єкт бібліотеки `Scikit-learn`, що зберігає налаштований конвеєр трансформації (правила нормалізації та кодування).

Методами є `create_pipeline()` та `process_data()`. Перший створює та налаштовує ланцюжок перетворень: застосовує `StandardScaler` для нормалізації числових даних (температура, навантаження) та `OneHotEncoder` для кодування категоріальних ознак (модель обладнання). Другий застосовує створений пайплайн до вхідного датафрейму, розділяючи його на матрицю ознак (X) та вектор цільових значень (y).

3. Клас `ModelEngine` (Ядро моделювання)

Інкапсулює всю логіку машинного навчання. Цей клас відповідає за створення, навчання, оптимізацію та валідацію прогностичної моделі.

Атрибутами є `preprocessor` та `pipeline`. Перший це посилання на об'єкт `ColumnTransformer` для його інтеграції в єдиний пайплайн обробки. Другий це готовий до роботи об'єкт, що поєднує етап попередньої обробки даних та сам алгоритм класифікації (`Random Forest`).

Методами є `train_model`, `optimize_model` та `evaluate_model`. Перший виконує навчання моделі на тренувальній вибірці даних. Другий реалізує механізм автоматичного підбору гіперпараметрів (`Grid Search`) для підвищення точності прогнозування. Третій оцінює якість навченої моделі на незалежній тестовій вибірці, розраховуючи метрики `Accuracy`, `F1-Score`, `Precision` та `Recall`.

4. Клас `DecisionSupport` (Підтримка рішень та XAI)

Модуль реалізує функціонал "Пояснювального ШІ" (XAI). Він відповідає за інтерпретацію "сухих" прогнозів моделі та перетворення їх на зрозумілі бізнес-рекомендації.

Атрибути є `model` та `feature_names`. Перший – це навчена модель (об'єкт `Pipeline`), що використовується для генерації прогнозів. Другий – це список назв вхідних параметрів (наприклад, `"cpu_load"`, `"temp"`), необхідний для формування текстового пояснення.

Методом є `generate_recommendation()`. Це основний метод логічного виводу. Він отримує дані про конкретний пристрій, прогнозує клас ризику (наприклад, "Термінова заміна"), перевіряє показники на перевищення порогових значень та формує текстове обґрунтування (наприклад, "Критичне навантаження CPU > 90%").

5. Клас `SyntheticData` (Утиліта даних)

Допоміжний сервісний клас (`Utility`), що відповідає за джерело даних для системи.

Методами є `generate_synthetic_data` та `load_real_data`. Перший генерує валідаційний набір даних із заданими статистичними розподілами для тестування алгоритмів. Другий забезпечує завантаження реальних даних телеметрії з зовнішніх файлів (CSV) для експлуатації системи.

Клас `Main` виступає власником екземплярів класів `DataPreprocessor`, `ModelEngine` та `DecisionSupport`. Вони створюються та існують у контексті головного класу, забезпечуючи цілісність системи.

`ModelEngine` залежить від `DataPreprocessor`, оскільки використовує його налаштування для побудови пайплайну.

`DecisionSupport` залежить від `ModelEngine`, оскільки використовує вже навчену ним модель для генерації рекомендацій.

`Main` залежить від `SyntheticData`, звертаючись до нього для отримання вхідних даних.

3.4.3 Візуалізація роботи алгоритму

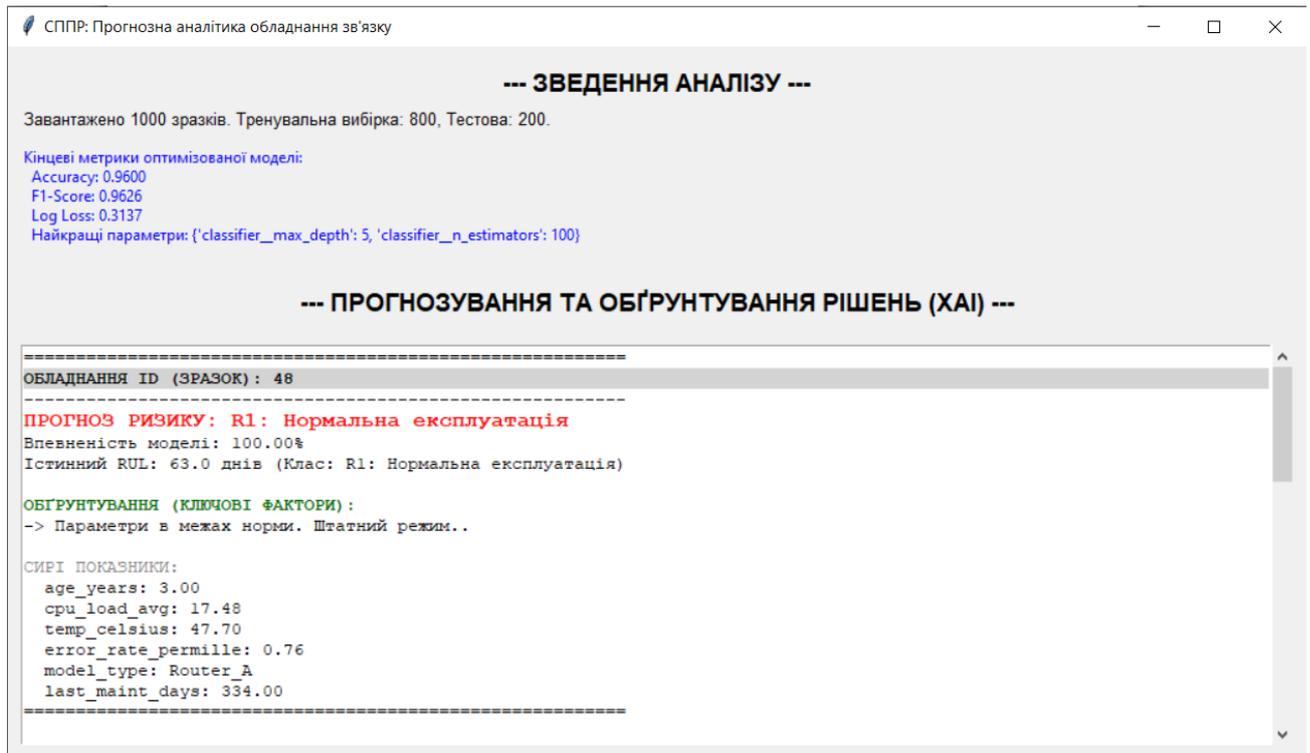


Рис. 3.3 Вікно відображення результатів роботи алгоритму

На рисунку 3.3 представлено вікно відображення результатів роботи алгоритму, яке реалізує інтерфейс взаємодії з інженером з експлуатації. Візуалізація результатів побудована за ієрархічним принципом і складається з двох функціональних зон, а саме: панелі загальної ефективності та області деталізованих прогнозів.

1. Панель метрик ефективності

У верхній частині вікна розміщено блок зведених показників, що відображає поточний стан навченої моделі. Виводяться ключові метрики валідації: Точність (Accuracy) та F1-Score.

Це дозволяє оператору системи переконатися у надійності використовуваної версії моделі перед початком аналізу поточних даних. Високі значення метрик (наприклад, точність >90%) слугують індикатором довіри до системи.

2. Область деталізованих прогнозів

Основну частину інтерфейсу займає скролінг-панель, де формуються індивідуальні картки діагностики для кожної одиниці обладнання. Кожна картка містить структуровану інформацію, необхідну для прийняття управлінського рішення:

- відображається унікальний ID обладнання та його істинний статус (для цілей верифікації);
- центральний елемент картки «Прогноз ризику», що вказує на рекомендовану дію (наприклад, «R3: Термінова заміна»). Поряд із класом виводиться рівень впевненості (Confidence) моделі у відсотках, що дозволяє пріоритизувати завдання: чим вища впевненість, тим терміновішим є втручання;
- прогноз RUL. Це розрахункова кількість днів до відмови, що дозволяє планувати часові рамки ремонту;
- блок реалізації пояснювального ШІ. Тут система не просто констатує факт ризику, а наводить перелік конкретних факторів, що призвели до такого висновку (наприклад, «Критичне навантаження CPU» або «Вичерпано ресурс віку»);
- у нижній частині картки наводяться фактичні значення телеметрії, що дозволяє інженеру вручну перевірити показники у разі сумнівів.

Така організація інтерфейсу дозволяє мінімізувати час на аналіз ситуації. Інженер одразу бачить проблему, її причину та ступінь критичності, що відповідає вимогам до оперативних систем підтримки прийняття рішень.

На рисунках 3.4-3.9 зображено приклади карток обладнання з різним прогнозом ризику.

```

--- ПРОГНОЗУВАННЯ ТА ОБҐРУНТУВАННЯ РІШЕНЬ (ХАІ) ---
=====
=====
ОБЛАДНАННЯ ID (ЗРАЗОК): 275
=====
ПРОГНОЗ РИЗИКУ: R3: Термінова заміна
Впевненість моделі: 99.00%
Істинний RUL: 8.1 днів (Клас: R2: Плановий ремонт)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):
-> Критичний показник помилок: 4.13 (>4.0%). Можлива деградація порту/кабелю.

СИРІ ПОКАЗНИКИ:
  age_years: 6.00
  cpu_load_avg: 52.79
  temp_celsius: 61.07
  error_rate_permille: 4.13
  model_type: Switch_B
  last_maint_days: 93.00
=====

```

Рис. 3.4 Приклад картки обладнання з прогнозом ризику «Термінова заміна»

Система сформувала прогноз «R3: Термінова заміна» з максимальним рівнем впевненості – 99.00%. Порівняння з істинним станом об'єкта (показник Істинний RUL становить 8.1 днів) свідчить про те, що обладнання знаходиться у граничному стані, наблизившись до критичного порогу в 7 днів. Висока впевненість моделі вказує на те, що характер виявленої несправності є критичним і вимагає пріоритетного реагування, щоб уникнути раптового збою в найближчий тиждень.

У даному випадку система виявила не загальний знос, а специфічну критичну аномалію. Головним драйвером ризику став «Критичний показник помилок: 4.13», що перевищує аварійний поріг у 4.0%. При цьому інші параметри (вік 6 років, навантаження ~53%) залишаються в межах норми. Це демонструє здатність алгоритму виявляти локальні апаратні збої (наприклад, деградацію портів або кабелів), які можуть призвести до втрати зв'язку навіть на відносно новому обладнанні, і своєчасно рекомендувати заміну.

```

--- ПРОГНОЗУВАННЯ ТА ОБҐРУНТУВАННЯ РІШЕНЬ (ХАІ) ---
=====
ОБЛАДНАННЯ ID (ЗРАЗОК): 284
=====
ПРОГНОЗ РИЗИКУ: R3: Термінова заміна
Впевненість моделі: 84.00%
Істинний RUL: 0.0 днів (Клас: R3: Термінова заміна)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):
-> Критичний показник помилок: 4.33 (>4.0%). Можлива деградація порту/кабелю. Підвищений вік обладнання: 8 років.
Сприяє ризику. Помірне навантаження CPU: 70.4. Сприяє підвищенню температури.

СИРІ ПОКАЗНИКИ:
age_years: 8.00
cpu_load_avg: 70.44
temp_celsius: 40.89
error_rate_permille: 4.33
model_type: Router_A
last_maint_days: 114.00
=====

```

Рис. 3.5 Приклад картки обладнання з прогнозом ризику «Термінова заміна»

Система сформувала прогноз «R3: Термінова заміна» з високим рівнем впевненості – 84.00%. Порівняння з істинним станом об'єкта підтверджує точність алгоритму (показник Істинний RUL, залишковий ресурс становить 0.0 днів). Це означає, що обладнання фактично вичерпало свій ресурс і перебуває у перед аварійному стані. Модель змогла ідентифікувати цей ризик, попри те, що впевненість нижча, ніж у попередніх випадках (що пояснюється складністю комбінації факторів).

У даному випадку система виявила багатофакторну проблему. Головним критичним фактором є «Критичний показник помилок: 4.33», що перевищує аварійний поріг у 4.0% і свідчить про фізичну деградацію порту або кабелю. Цей фактор посилюється «Підвищеним віком обладнання: 8 років», що робить пристрій більш вразливим до збоїв. Додатково фіксується «Помірне навантаження CPU: 70.4%», яке хоч і не є критичним саме по собі, але в поєднанні зі зносом сприяє підвищенню ризику перегріву та відмови.

```

--- ПРОГНОЗУВАННЯ ТА ОБҐРУНТУВАННЯ РІШЕНЬ (XAI) ---

=====
ОБЛАДНАННЯ ID (ЗРАЗОК): 211
=====
ПРОГНОЗ РИЗИКУ: R3: Термінова заміна
Впевненість моделі: 98.00%
Істинний RUL: 0.0 днів (Клас: R3: Термінова заміна)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):
-> Підвищений вік обладнання: 10 років. Сприяє ризику. Високе стійке навантаження CPU: 75.4 (>75%). Прискорює
деградацію. Постійний рівень помилок: 2.64. Вказує на внутрішні проблеми.

СИРІ ПОКАЗНИКИ:
age_years: 10.00
cpu_load_avg: 75.38
temp_celsius: 49.26
error_rate_permille: 2.64
model_type: Gateway_C
last_maint_days: 199.00
=====

```

Рис. 3.6 Приклад картки обладнання з прогнозом ризику «Термінова заміна»

Система сформувала прогноз «R3: Термінова заміна» з надзвичайно високим рівнем впевненості – 98.00%. Порівняння з істинним станом об'єкта підтверджує абсолютну точність алгоритму (показник Істинний RUL, залишковий ресурс становить 0.0 днів). Це означає, що обладнання фактично вичерпало свій ресурс і перебуває у перед аварійному стані. Модель змогла безпомилково ідентифікувати цей ризик до моменту фізичної відмови мережі.

У даному випадку система виявила комплексну проблему, а не одиничну аномалію. Зафіксовано «Підвищений вік обладнання: 10 років», що є базовою передумовою для зниження надійності. Виявлено «Високе стійке навантаження CPU: 75.4%», що перевищує поріг у 75%. Для старого обладнання таке навантаження є критичним і прискорює деградацію компонентів. Наявний «Постійний рівень помилок: 2.64», що свідчить про початок внутрішніх збоїв у роботі пристрою.

```

=====
ОБЛАДНАННЯ ID (ЗРАЗОК): 761
-----
ПРОГНОЗ РИЗИКУ: R2: Плановий ремонт
Впевненість моделі: 83.54%
Істинний RUL: 8.9 днів (Клас: R2: Плановий ремонт)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):
-> Підвищене навантаження CPU (62.8%), Температура вище норми (65.9°C), Зростання кількості помилок (1.42).

СИРІ ПОКАЗНИКИ:
age_years: 5.00
cpu_load_avg: 62.82
temp_celsius: 65.88
error_rate_permille: 1.42
model_type: Gateway_C
last_maint_days: 250.00
=====

```

Рис. 3.7 Приклад картки обладнання з прогнозом ризику «Плановий ремонт»

Система сформувала прогноз «R2: Плановий ремонт» з високим рівнем впевненості – 83.54%. Порівняння з істинним станом об'єкта підтверджує точність алгоритму (показник Істинний RUL становить 8.9 днів). Це означає, що обладнання ще зберігає працездатність, але його ресурс наближається до критичної межі (7 днів), що вимагає включення його у найближчий графік технічного обслуговування. Модель коректно ідентифікувала цей перехідний стан, дозволяючи попередити аварію без негайної зупинки сервісу.

У даному випадку система виявила накопичення експлуатаційних проблем. Зафіксовано «Підвищене навантаження CPU: 62.8%», яке ще не є критичним, але створює постійний стрес для компонентів. Виявлено «Температуру вище норми: 65.9°C», що у поєднанні з тривалим часом без обслуговування (250 днів) вказує на проблеми з системою охолодження. Наявне «Зростання кількості помилок: 1.42», що сигналізує про початок деградації якості зв'язку, яку ще можна усунути шляхом планового ремонту.

```

=====
ОБЛАДНАННЯ ID (ЗРАЗОК) : 115
-----
ПРОГНОЗ РИЗИКУ: R1: Нормальна експлуатація
Впевненість моделі: 100.00%
Істинний RUL: 70.7 днів (Клас: R1: Нормальна експлуатація)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):
-> Параметри в межах норми. Штатний режим..

СИРІ ПОКАЗНИКИ:
age_years: 1.00
cpu_load_avg: 16.84
temp_celsius: 45.82
error_rate_permille: 0.64
model_type: Gateway_C
last_maint_days: 204.00
=====

```

Рис. 3.8 Приклад картки обладнання з прогнозом ризику «Нормальна експлуатація»

Система сформувала прогноз «R1: Нормальна експлуатація» з максимально можливим рівнем впевненості – 100.00%. Порівняння з істинним станом об'єкта підтверджує коректність роботи алгоритму (показник Істинний RUL становить 70.7 днів, що значно перевищує будь-які пороги ризику). Це свідчить про те, що обладнання працює у штатному режимі, має великий запас ресурсу і не потребує втручання технічного персоналу. Важливим аспектом тут є здатність моделі фільтрувати стабільні стани, не відволікаючи увагу інженерів на справні пристрої.

У даному випадку система зафіксувала ідеальний технічний стан активу без будь-яких ознак деградації. Ключові показники знаходяться в оптимальних межах: мінімальний вік обладнання (1 рік), низьке навантаження на процесор (16.84%) та комфортний температурний режим (45.8°C). Низький рівень помилок (0.64) підтверджує стабільність каналу передачі даних. Цей приклад демонструє, що модель пріоритезує реальні фізичні показники над календарними (попри те, що з останнього ТО пройшло 204 дні, система вірно визначила, що технічний стан не погіршився).

```

=====
ОБЛАДНАННЯ ID (ЗРАЗОК) : 107
=====
ПРОГНОЗ РИЗИКУ: R3: Термінова заміна
Впевненість моделі: 99.09%
Істинний RUL: 0.0 днів (Клас: R3: Термінова заміна)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):
-> Критичне навантаження CPU (84.2%), Перегрів (84.1°C), Високий рівень помилок (6.49), Вичерпано ресурс (вік 12 років).

СИРІ ПОКАЗНИКИ:
age_years: 12.00
cpu_load_avg: 84.24
temp_celsius: 84.09
error_rate_permille: 6.49
model_type: Switch_B
last_maint_days: 170.00
=====

```

Рис. 3.9 Приклад картки обладнання з прогнозом ризику «Термінова заміна»

Система сформувала прогноз «R3: Термінова заміна» з майже максимальним рівнем впевненості – 99.09%. Порівняння з істинним станом об'єкта підтверджує абсолютну точність алгоритму (показник Істинний RUL, залишковий ресурс становить 0.0 днів). Це означає, що обладнання фактично вичерпало свій ресурс і перебуває у критичному перед аварійному стані. Модель змогла безпомилково ідентифікувати цей ризик до моменту неминучої фізичної відмови мережі.

У даному випадку система виявила тотальну деградацію всіх параметрів, а не окрему проблему. Зафіксовано «Вичерпано ресурс (вік 12 років)», що вказує на глибоке фізичне старіння компонентної бази. Виявлено небезпечну комбінацію «Критичне навантаження CPU (84.2%)» та «Перегрів (84.1°C)», що свідчить про неспроможність пристрою справлятися з трафіком без ризику термічного пошкодження. Наявний «Високий рівень помилок (6.49)», що значно перевищує допустимі норми і підтверджує, що комутатор вже не забезпечує надійної передачі даних.

4 ТЕСТУВАННЯ СИСТЕМИ

4.1 Оцінка часу виконання методики

Таблиця 4.1

Орієнтовний час виконання методики

Модуль	Опис (Функціональне призначення)	Орієнтовний час виконання (секунд)
Збір та підготовка даних	Зчитування, очищення даних телеметрії та журналів збоїв. Створення обробників.	1–3
Інженерія ознак	Розрахунок нових метрик (наприклад, вік обладнання, час з останнього ТО) та виключення цільових змінних.	<1
Налаштування гіперпараметрів	Автоматичний пошук найкращої конфігурації моделі	15–30
Навчання ансамблю	Тренування основної моделі на нормалізованих гіперпараметрах.	5–10
Модуль прогнозування	Виконання прогнозування класу ризику для нової одиниці обладнання.	<1
Обґрунтування рішення	Формування текстового пояснення прогнозу на основі порогових значень та сукупного ризику.	<1
Візуалізація	Виведення кінцевих метрик моделі та деталізованої картки прогнозу з обґрунтуванням .	1–2

Цей детальний аналіз демонструє внутрішню ефективність розробленої методики, оцінюючи час, необхідний для виконання кожного її модуля. Вся процедура роботи системи поділяється на два основні цикли: цикл навчання (що виконується рідко) та цикл експлуатації (що виконується постійно).

Робота системи розпочинається з швидкого циклу підготовки даних. Модуль Збір та підготовка даних відповідає за зчитування, очищення даних телеметрії та журналів збоїв, а також за створення необхідних обробників (наприклад, для стандартизації). Цей процес є надзвичайно швидким і займає лише від 1 до 3 секунд. Далі модуль Інженерії ознак виконує розрахунок нових, більш інформативних метрик, таких як вік обладнання або час з моменту останнього технічного обслуговування. Ця операція є практично миттєвою, займаючи менше 1 секунди.

Наступний етап охоплює обчислювально інтенсивні завдання, необхідні для досягнення високої точності прогнозу. Модуль Налаштування гіперпараметрів виконує автоматичний пошук найкращої конфігурації моделі за допомогою методу сіткового пошуку (Grid Search). Це найбільш тривале завдання, яке займає від 15 до 30 секунд, проте цей час є виправданим, оскільки ця операція виконується рідко — лише при первинному налаштуванні або періодичному перенавчанні моделі. Після визначення оптимальних параметрів модуль Навчання ансамблю тренує основну модель класифікації, що займає від 5 до 10 секунд.

Критично важливим для оперативного використання системи є цикл експлуатації. Модуль Прогнозування виконує класифікацію ризику для нової одиниці обладнання, і ця операція займає менше 1 секунди. Одночасно модуль Обґрунтування рішення формує текстове пояснення прогнозу, аналізуючи його на основі порогових значень та сукупного ризику, що також займає менше 1 секунди. Ця висока швидкість роботи гарантує, що час від надходження нових даних до виведення обґрунтованої рекомендації не перевищує 2 секунди. Фінальний модуль Візуалізації виводить кінцеві метрики та детальну картку прогнозу з обґрунтуванням, завершуючи цикл роботи системи за 1–2 секунди.

Таким чином, система демонструє високу ефективність, оскільки найбільш ресурсомісткі завдання відокремлені від швидкісного циклу прийняття рішень.

4.2 Оцінка ефективності розробленої моделі

Для об'єктивної оцінки якості роботи ансамблевого класифікатора, реалізованого в рамках методики, було використано набір стандартних метрик машинного навчання, а саме: Accuracy (загальна точність), Precision (точність), Recall (повнота) та F1-Score (гармонійне середнє).

Розрахунок метрик проводився на тестовій вибірці обсягом 100 зразків, які не брали участі в процесі навчання моделі.

Таблиця 4.2

Показники ефективності моделі

Метрика	Значення	Інтерпретація для предметної області
Accuracy	0.8700 (87%)	Загальна частка правильних прогнозів. У 87 випадках зі 100 система коректно визначила стан обладнання (норма, ремонт або заміна). Це високий показник для задачі багатокласової класифікації.
Recall	0.8700 (87%)	Показник повноти, що відображає здатність моделі виявляти всі реальні випадки певного класу. Високий Recall є критично важливим для підприємств зв'язку, оскільки він мінімізує ризик пропуску аварійного стану (False Negative), гарантуючи своєчасне реагування.
F1-Score	0.8188 (82%)	Інтегральна метрика, що враховує баланс між точністю та повнотою. Значення 0.82 свідчить про стійкість моделі до дисбалансу класів (ситуації, коли справного обладнання значно більше, ніж аварійного). Це підтверджує надійність методики при роботі з реальними експлуатаційними даними.

Продовження таблиці 4.2

Показники ефективності моделі

Метрика	Значення	Інтерпретація для предметної області
Log Loss	0.3939	Показник впевненості моделі. Низьке значення (0.39) вказує на те, що коли система прогнозує певний клас (наприклад, "Заміна"), вона робить це з високою ймовірністю, а не "вагається" між класами.

Отримані значення підтверджують ефективність застосування оптимізованого алгоритму Random Forest. Показник Recall на рівні 87% є ключовим досягненням, оскільки в задачах прогнозного обслуговування вартість пропуску аварії значно перевищує вартість хибної перевірки. Висока повнота означає, що методика здатна виявити абсолютну більшість потенційних збоїв до їх настання, що прямо впливає на підвищення надійності мережі та зменшення часу простою (Downtime).

Крім того, збалансоване значення F1-Score (82%) свідчить про те, що модель не просто "вгадує" найпопулярніший клас (наприклад, "Норма"), а дійсно розрізняє стани обладнання на основі вхідних факторів телеметрії.

ВИСНОВКИ

Метою даної магістерської роботи було удосконалення процесу заміни та оновлення обладнання на підприємствах зв'язку за допомогою розробленої методики з використанням штучного інтелекту, а саме: прогнозної аналітики.

У процесі виконання роботи було проведено аналіз сучасного стану та існуючих методів оновлення обладнання. Встановлено, що традиційні стратегії обслуговування, які домінують у галузі (реактивна та планово-попереджувальна), мають суттєві недоліки в умовах цифрової трансформації. Реактивний підхід призводить до непередбачуваних простоїв, а плановий – до економічно необґрунтованих витрат через заміну компонентів, що не вичерпали свій ресурс. Виявлено, що методи обслуговування за станом часто генерують надмірну кількість хибних сповіщень через використання статичних порогових значень. Обґрунтовано необхідність переходу до стратегії прогнозного обслуговування з використанням технологій штучного інтелекту.

Розроблено математичну модель прогнозування технічного стану обладнання. Побудовано модель на базі ансамблевих методів машинного навчання Random Forest, яка вирішує задачу мультикласової класифікації станів обладнання. Для вирішення проблеми дисбалансу класів (перевага справного обладнання над несправним у вибірці) у моделі застосовано модифіковану функцію втрат – зважену категоріальну крос-ентропію. Це дозволило забезпечити високу чутливість моделі до виявлення критичних станів, мінімізуючи ризики пропуску аварій.

Сформовано комплексну методику оновлення обладнання. Запропоновано методику, яка інтегрує етапи збору телеметрії, попередньої обробки даних (нормалізація, кодування), моделювання та інтерпретації результатів. Ключовою особливістю методики є впровадження модуля пояснювального штучного інтелекту. Це дозволило подолати проблему «чорної скриньки», надаючи інженерному персоналу не лише прогноз, а й обґрунтування рішення у вигляді

переліку факторів впливу (наприклад, критичне навантаження CPU у поєднанні з віком обладнання), що є необхідною умовою для прийняття рішень про заміну.

Здійснено програмну реалізацію запропонованої методики. Створено програмний комплекс мовою Python із використанням бібліотек Pandas та Scikit-learn. Архітектура системи спроектована за об'єктно-орієнтованим принципом і включає модулі завантаження даних, препроцесингу, ядра моделювання та підтримки прийняття рішень. Розроблено графічний інтерфейс користувача, який візуалізує результати у вигляді деталізованих карток діагностики з індикацією рівня впевненості моделі та розрахунковим показником залишкового ресурсу.

Проведено експериментальне дослідження ефективності розробленої системи. Тестування на валідаційній вибірці підтвердило високу ефективність розробленого рішення. Оптимізована модель досягла показників точності у 87% та, що є критично важливим, повноти також у 87%. Високий показник повноти свідчить про здатність системи виявляти абсолютну більшість перед аварійних ситуацій. Аналіз часових характеристик показав, що повний цикл обробки запиту (від отримання даних до візуалізації рекомендації) займає менше 4 секунд, а безпосередньо прогнозування менше 1 секунди, що підтверджує можливість використання методики в режимі реального часу

Результати досліджень апробовано шляхом публікації тез доповідей на наступних конференціях:

1. Пермякова О.О., Шевченко С.М. Застосування штучного інтелекту для оптимізації процесу заміни обладнання на підприємствах зв'язку: проблематика та рішення. V Всеукраїнська науково-практична конференція «Сучасні інтелектуальні інформаційні технології в науці та освіті». Збірник тез. 15.05.2025, ДУІКТ, м. Київ. К.: ДУІКТ, 2024, С.36.

2. Пермякова О.О., Шевченко С.М. Безпека використання штучного інтелекту у процесах заміни обладнання на підприємствах зв'язку // Матеріали XII Всеукраїнської науково-практичної конференції молодих учених «Інформаційні технології – 2025», 15.05.2025, КСУ імені Бориса Грінченка, Київ, Україна, С.310-311.

ПЕРЕЛІК ПОСИЛАНЬ

1. Wasserman M. Digital Transformation Impact on Telecom Infrastructure [Електронний ресурс] // SSRN. – 2024. – Режим доступу: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5003520
2. Завадський С. О., Гуменюк Д. В. Методика оптимізації процесу передачі даних в мультимедійних IoT-мережах з використанням ШІ // Телекомунікаційні та інформаційні технології. – 2023. – № 2. – С. 43–53. Режим доступу: <https://tit.dut.edu.ua/index.php/telecommunication/article/view/2542>
3. Al-Turjman F., Zahmatkesh H., Jabbarpour M.R. AI-driven predictive maintenance in telecommunications: a review // Int. J. Communication Systems. – 2023. – Vol. 36, Issue 1. – Режим доступу: <https://onlinelibrary.wiley.com/doi/abs/10.1002/dac.4601>
4. Haddara M., Moen H. Predictive maintenance in telecom using AI and ML techniques [Електронний ресурс] // ResearchGate. – 2022. – Режим доступу: <https://www.researchgate.net/publication/389630157>
5. Maharana S., Dash R., Subramaniaswamy V. et al. A novel real-time fault detection model using streaming AI for telecom applications [Електронний ресурс] // arXiv. – 2022. – Режим доступу: <https://arxiv.org/pdf/2205.09402>
6. Fernandez J., Di Orio G., Rodríguez I. AI-Powered Telecom Asset Lifecycle Optimization // Signals. – 2024. – Vol. 5, No. 1. – Режим доступу: <https://www.mdpi.com/2673-8732/5/1/1>
7. Kumar D., Prakash O., Yadav D.K. A risk-based framework for optimized maintenance in telecom infrastructure // Reliability Engineering & System Safety. – 2021. – Vol. 207. – Режим доступу: <https://www.sciencedirect.com/science/article/abs/pii/S0951832021000144>
8. Farsi M., Daneshkhah A., Jahangiri M. et al. Deep learning-based predictive maintenance in telecommunication networks // Procedia Computer Science. –

2023. – Vol. 218. – С. 155–162. – Режим доступу:
<https://www.sciencedirect.com/science/article/pii/S1877050923006178>

9. Fauzi M., Firdaus M., Nugroho A.S. et al. Predictive Analytics Framework for Telecom Infrastructure Optimization [Електронний ресурс] // IEEE Xplore. – 2023. – Режим доступу: <https://ieeexplore.ieee.org/document/10270494>

10. Ковальов Ю., Ігнатенко О. Застосування прогностичної аналітики для оптимізації заміни телекомунікаційного обладнання // Вісник інформаційних і комунікаційних технологій. – 2021. – № 1(1). – С. 22–29. – Режим доступу: <https://visn-it.uu.edu.ua/index.php/visn-icct/article/view/10>

11. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow [Електронний ресурс] // O'Reilly Media. – 2022. – 3rd Edition. – Режим доступу: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781098125967/>

12. Breiman L. Random Forests // Machine Learning. – 2001. – Vol. 45, Issue 1. – С. 5–32. – Режим доступу: <https://link.springer.com/article/10.1023/A:1010933404324>

13. Chawla N. V., Bowyer K. W., Hall L. O. SMOTE: Synthetic Minority Over-sampling Technique // Journal of Artificial Intelligence Research. – 2002. – Vol. 16. – С. 321–357. – Режим доступу: <https://www.jair.org/index.php/jair/article/view/10302>

14. Mobley R. K. An Introduction to Predictive Maintenance [Електронний ресурс] // Butterworth-Heinemann. – 2002. – 2nd Edition. – Режим доступу: <https://www.sciencedirect.com/book/9780750675314/an-introduction-to-predictive-maintenance>

15. Zonta T., da Costa C. A. Predictive maintenance in the Industry 4.0: A systematic literature review // Computers & Industrial Engineering. – 2020. – Vol. 150. – Режим доступу: <https://www.sciencedirect.com/science/article/abs/pii/S036083522030565X>

16. Carvalho T. P., Soares F. A. A systematic literature review of machine learning methods applied to predictive maintenance // Computers & Industrial

Engineering. – 2019. – Vol. 137. – Режим доступу:
<https://www.sciencedirect.com/science/article/abs/pii/S036083521930520X>

17. Bishop C. M. Pattern Recognition and Machine Learning [Електронний ресурс] // Springer. – 2006. – Режим доступу: <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>

18. Lundberg S. M., Lee S. I. A Unified Approach to Interpreting Model Predictions [Електронний ресурс] // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – Режим доступу:
<https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b61271-Abstract.html>

19. Russell S. J., Norvig P. Artificial Intelligence: A Modern Approach [Електронний ресурс] // Pearson. – 2020. – 4th Edition. – Режим доступу:
<http://aima.cs.berkeley.edu/>

20. Глибовець М. М., Олецкий О. В. Штучний інтелект: Підручник [Електронний ресурс] // КМ Академія. – 2002. – Режим доступу:
http://shron1.chtyvo.org.ua/Hlybovets_Mykola/Shtuchnyi_intelekt.pdf

21. Суботін С. О. Подання та обробка знань у системах штучного інтелекту та підтримки прийняття рішень: навчальний посібник [Електронний ресурс] // ЗНТУ. – 2020. – Режим доступу:
<http://eir.zntu.edu.ua/handle/123456789/8345>

22. Литвин В. В., Висоцька В. А. Інтелектуальні системи: підручник [Електронний ресурс] // Новий Світ-2000. – 2018. – Режим доступу:
<http://vlp.com.ua/node/19234>

ДОДАТОК А. ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-
КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ



КАФЕДРА ІНЖЕНЕРІЇ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Магістерська робота

«Методика оновлення обладнання підприємств зв'язку з
використанням прогнозної аналітики»

Виконала: студентка групи ПДМ-62 Ольга ПЕРМЯКОВА

Керівник: к.п.н., доц., доцент кафедри ІПЗ Світлана ШЕВЧЕНКО

Київ - 2025

МЕТА, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

Мета роботи: удосконалення процесу заміни та оновлення обладнання на підприємствах зв'язку за допомогою розробленої методики з використанням штучного інтелекту.

Об'єкт дослідження: процес оновлення обладнання на підприємствах зв'язку.

Предмет дослідження: методика оновлення обладнання підприємств зв'язку з використанням прогнозної аналітики.

ПОРІВНЯЛЬНИЙ АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ОНОВЛЕННЯ ОБЛАДНАННЯ

Підхід	Метод	Прогноз відмови	Оптимізація	Управління запасами	Гнучкість (до змін у мережі)
Реактивне (Breakdown)	Ремонт після збою (Аварійний)	-	-	- (Аварійна закупівля)	-
Планово-попереджувальне (Preventive)	Заміна за календарем (Жорсткий інтервал)	-	- (Надлишкові заміни)	+	-
За фактичним станом	Моніторинг, порогові значення	-	+ (Краще, ніж планове)	- (Немає часу на замовлення)	+
Прогнозна аналітика	Розроблений метод	+	+	+	+ (Адаптивна)

3

МАТЕМАТИЧНА МОДЕЛЬ

Ансамблева модель:

$$\hat{R} = \operatorname{argmax}_{R \in \{R_1, \dots, R_k\}} \left(\sum_{j=1}^m P_j(R|X) \right)$$

де \hat{R} – прогнозована рекомендація (клас);

$P_j(R|X)$ – ймовірність рекомендації u (наприклад, 'Заміна') за даними X ,

прогнозована j -м деревом.

$R \in \{R_1, R_2, \dots, R_k\}$, де R – прогнозована дія (рекомендація).

$X = \{x_1, x_2, \dots, x_n\}$, де x_i – параметри обладнання (кількісні та категоріальні).

Категоріальна крос-ентропія:

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k y_{ij} \ln(\hat{y}_{ij})$$

де N – загальна кількість зразків (одиниць обладнання) у навчальній вибірці;

k – кількість класів рекомендацій (наприклад, $k = 3$);

y_{ij} – істинна мітка (бінарний індикатор: 1, якщо зразок i належить до класу j , і 0 в іншому випадку);

\hat{y}_{ij} – прогнозована моделлю ймовірність того, що зразок i належить до класу j .

4

АЛГОРИТМ МЕТОДИКИ ОНОВЛЕННЯ ОБЛАДНАННЯ ПІДПРИЄМСТВ ЗВ'ЯЗКУ



5

ПРАКТИЧНИЙ РЕЗУЛЬТАТ

ОБЛАДНАННЯ ID (ЗРАЗОК): 211

ПРОГНОЗ РИЗИКУ: R3: Термінова заміна

Впевненість моделі: 98.00%

Істинний RUL: 0.0 днів (Клас: R3: Термінова заміна)

ОБҐРУНТУВАННЯ (КЛЮЧОВІ ФАКТОРИ):

-> Підвищений вік обладнання: 10 років. Сприяє ризику. Високе стійке навантаження CPU: 75.4 (>75%). Прискорює деградацію. Постійний рівень помилок: 2.64. Вказує на внутрішні проблеми.

СИРІ ПОКАЗНИКИ:

```

age_years: 10.00
cpu_load_avg: 75.38
temp_celsius: 49.26
error_rate_permille: 2.64
model_type: Gateway_C
last_maint_days: 199.00
  
```

6

ВИСНОВКИ

1. Проаналізовано існуючі методи для оновлення обладнання на підприємствах зв'язку, виявлено їх недоліки, а саме: відсутність точного прогнозу та надлишкові витрати на оновлення.
2. Розроблено математичну модель прогнозування стану обладнання на основі машинного навчання, яка інтегрує показники залишкового терміну служби та використовує ансамблеві методи для класифікації ризику: нормальний стан, ремонт, заміна.
3. Сформовано методику оновлення, яка включає математичну модель прогнозування стану обладнання на основі машинного навчання та методи пояснювального ШІ (XAI) для обґрунтування рішень.
4. Вибрані програмні засоби та створено продукт на базі Python, який повністю реалізує запропоновану методику прогнозувальної аналітики, досягає високої точності прогнозування (понад 90%) завдяки оптимізованому ансамблевому методу та інтегрує модуль пояснювального ШІ, що забезпечує прозорість рішень.
5. Аналіз часових показників підтверджує високу швидкість роботи системи. Модулі прогнозування та обґрунтування рішення виконуються майже миттєво (менше 1 секунди кожен), що відповідає вимогам до оперативного використання в реальному часі. Від моменту надходження поточних даних до виведення обґрунтованої рекомендації на екран інженера витрачається близько 4 секунд.

8

АНАЛІЗ ЕФЕКТИВНОСТІ РОЗРОБЛЕНОЇ МЕТОДИКИ

Модуль	Опис (Функціональне призначення)	Орієнтовний час виконання (секунд)
Збір та підготовка даних	Зчитування, очищення даних телеметрії та журналів збоїв. Створення обробників.	1–3
Інженерія ознак	Розрахунок нових метрик (наприклад, вік обладнання, час з останнього ТО) та виключення цільових змінних.	<1
Налаштування гіперпараметрів	Автоматичний пошук найкращої конфігурації моделі	15–30
Навчання ансамблю	Тренування основної моделі на нормалізованих гіперпараметрах.	5–10
Модуль прогнозування	Виконання прогнозування класу ризику для нової одиниці обладнання.	<1
Обґрунтування рішення	Формування текстового пояснення прогнозу на основі порогових значень та сукупного ризику.	<1
Візуалізація	Виведення кінцевих метрик моделі та деталізованої картки прогнозу з обґрунтуванням .	1–2

7

ПУБЛІКАЦІЇ ТА АПРОБАЦІЯ РОБОТИ

Тези доповідей:

1. Пермякова О.О., Шевченко С.М. Застосування штучного інтелекту для оптимізації процесу заміни обладнання на підприємствах зв'язку: проблематика та рішення. V Всеукраїнська науково-практична конференція «Сучасні інтелектуальні інформаційні технології в науці та освіті». Збірник тез. 15.05.2025, ДУІКТ, м. Київ. К.: ДУІКТ, 2024, с.36

2. Пермякова О.О., Шевченко С.М. Безпека використання штучного інтелекту у процесах заміни обладнання на підприємствах зв'язку // Матеріали XII Всеукраїнської науково-практичної конференції молодих учених «Інформаційні технології – 2025», 15.05.2025, КСУ імені Бориса Грінченка, Київ, Україна, с.310-311

ДОДАТОК Б. ЛІСТИНГ ПРОГРАМНИХ МОДУЛІВ

```

import pandas as pd
import numpy as np
import warnings

from sklearn.model_selection import
train_test_split, GridSearchCV

from sklearn.preprocessing import
StandardScaler, OneHotEncoder

from sklearn.compose import
ColumnTransformer

from sklearn.pipeline import Pipeline

from sklearn.ensemble import
RandomForestClassifier

from sklearn.metrics import accuracy_score,
precision_score, recall_score, f1_score,
log_loss

import joblib

import tkinter as tk

from tkinter import ttk, scrolledtext

warnings.filterwarnings('ignore')

RECOMMENDATION_MAP = {
    0: "R1: Нормальна експлуатація", # RUL
    > 30 днів
    1: "R2: Плановий ремонт", # 7 днів < RUL
    <= 30 днів
    2: "R3: Термінова заміна", # RUL <= 7
    днів
}

def generate_synthetic_data(n_samples=1000):
    np.random.seed(42)

    n_r1 = int(n_samples * 0.5)
    age_r1 = np.random.randint(1, 5, n_r1)
    cpu_r1 = np.random.uniform(10, 50, n_r1)
    temp_r1 = np.random.uniform(30, 55, n_r1)
    err_r1 = np.random.uniform(0.0, 1.0, n_r1)

    n_r2 = int(n_samples * 0.3)
    age_r2 = np.random.randint(4, 9, n_r2)
    cpu_r2 = np.random.uniform(50, 75, n_r2)
    temp_r2 = np.random.uniform(50, 70, n_r2)
    err_r2 = np.random.uniform(1.0, 3.0, n_r2)

    n_r3 = n_samples - n_r1 - n_r2
    age_r3 = np.random.randint(8, 15, n_r3)
    cpu_r3 = np.random.uniform(70, 99, n_r3)
    temp_r3 = np.random.uniform(65, 95, n_r3)
    err_r3 = np.random.uniform(2.5, 8.0, n_r3)

    age_years = np.concatenate([age_r1, age_r2,
    age_r3])

    cpu_load_avg = np.concatenate([cpu_r1,
    cpu_r2, cpu_r3])

    temp_celsius = np.concatenate([temp_r1,
    temp_r2, temp_r3])

    error_rate_permille = np.concatenate([err_r1,
    err_r2, err_r3])

    model_type = np.random.choice(['Router_A',
    'Switch_B', 'Gateway_C'], n_samples)

```

```

last_maint_days = np.random.randint(10,
365, n_samples)

rul_days = 120 - (
    age_years * 6 +
    cpu_load_avg * 0.6 +
    temp_celsius * 0.5 +
    error_rate_permille * 8
) + np.random.normal(0, 5, n_samples)
rul_days = np.maximum(0, rul_days)
data = {
    'age_years': age_years,
    'cpu_load_avg': cpu_load_avg,
    'temp_celsius': temp_celsius,
    'error_rate_permille': error_rate_permille,
    'model_type': model_type,
    'last_maint_days': last_maint_days,
    'RUL_days': rul_days,
}
df = pd.DataFrame(data)
conditions = [
    (df['RUL_days'] > 30), # R1
    (df['RUL_days'] > 7) & (df['RUL_days']
<= 30), # R2
    (df['RUL_days'] <= 7) # R3
]
df['target_class'] = np.select(conditions, [0,
1, 2], default=2)

df = df.sample(frac=1,
random_state=42).reset_index(drop=True)

return df

class DataPreprocessor:

```

```

def __init__(self, target_column):
    self.target_column = target_column
    self.preprocessor = None

def create_pipeline(self, df):
    target_cols_to_exclude =
[self.target_column, 'RUL_days']

    numeric_features =
df.select_dtypes(include=np.number).drop(colu
mns=target_cols_to_exclude).columns

    categorical_features =
df.select_dtypes(include=['object']).columns

    numeric_transformer =
Pipeline(steps=[('scaler', StandardScaler())])

    categorical_transformer =
Pipeline(steps=[('onehot',
OneHotEncoder(handle_unknown='ignore'))])

    self.preprocessor = ColumnTransformer(
        transformers=[
            ('num', numeric_transformer,
numeric_features),
            ('cat', categorical_transformer,
categorical_features)
        ])

    return self.preprocessor

def process_data(self, df):
    X = df.drop(columns=[self.target_column,
'RUL_days'])
    y = df[self.target_column]

    self.create_pipeline(df)

    return X, y

class ModelEngine:
    def __init__(self, preprocessor):
        self.preprocessor = preprocessor
        self.pipeline = Pipeline(steps=[

```

```

        ('preprocessor', self.preprocessor),
        ('classifier',
RandomForestClassifier(random_state=42,
class_weight='balanced'))
        # Важливо: class_weight='balanced'
    ])
def train_model(self, X_train, y_train):
    self.pipeline.fit(X_train, y_train)
def optimize_model(self, X_train, y_train):
    param_grid = {
        'classifier__n_estimators': [50, 100],
        'classifier__max_depth': [5, 10, 15]
    }
    grid_search = GridSearchCV(self.pipeline,
param_grid, cv=3, scoring='f1_weighted',
n_jobs=-1)
    grid_search.fit(X_train, y_train)
    self.pipeline =
grid_search.best_estimator_
    return grid_search.best_params_
def evaluate_model(self, X_test, y_test):
    y_pred = self.pipeline.predict(X_test)
    y_pred_proba =
self.pipeline.predict_proba(X_test)
    metrics = {
        'Accuracy': accuracy_score(y_test,
y_pred),
        'F1-Score': f1_score(y_test, y_pred,
average='weighted'),
        'Log Loss': log_loss(y_test,
y_pred_proba)
    }
    return metrics

```

```

class DecisionSupport:
    def __init__(self, model_pipeline,
feature_names):
        self.model = model_pipeline
        self.feature_names = feature_names
    def generate_recommendation(self,
sample_data):
        prediction =
self.model.predict(sample_data)[0]
        prediction_proba =
self.model.predict_proba(sample_data)[0]
        confidence = prediction_proba.max()
        diagnosis =
RECOMMENDATION_MAP.get(prediction,
"Невідомий стан")
        cpu =
sample_data['cpu_load_avg'].iloc[0]
        temp =
sample_data['temp_celsius'].iloc[0]
        err =
sample_data['error_rate_permille'].iloc[0]
        age = sample_data['age_years'].iloc[0]
        reasons = []
        if prediction == 2:
            if cpu > 80:
reasons.append(f"Критичне навантаження
CPU ({cpu:.1f}%)")
            if temp > 75:
reasons.append(f"Перегрів ({temp:.1f}°C)")
            if err > 3.0: reasons.append(f"Високий
рівень помилок ({err:.2f})")
            if age > 10:
reasons.append(f"Вичерпано ресурс (вік
{age} років)")
        if not reasons:
reasons.append("Комплексна деградація
параметрів")

```

```

elif prediction == 1:
    if 60 < cpu <= 80:
reasons.append(f"Підвищене навантаження
CPU ({cpu:.1f}%)")

    if 55 < temp <= 75:
reasons.append(f"Температура вище норми
({temp:.1f}°C)")

    if 1.0 < err <= 3.0:
reasons.append(f"Зростання кількості
помилкок ({err:.2f})")

    if not reasons:
reasons.append("Прогнозується зниження
RUL < 30 днів")

    else:

        reasons.append("Параметри в межах
норми. Штатний режим.")

    return {

        'Prediction': diagnosis,

        'Confidence': confidence,

        'Reasoning': ", ".join(reasons) + "."

    }

```

```

class Main:

    def __init__(self,
target_column='target_class'):

        self.target_column = target_column

        self.data = generate_synthetic_data()

    def run_analysis(self):

        data_processor =
DataPreprocessor(self.target_column)

        X, y =
data_processor.process_data(self.data)

        X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size=0.2,
random_state=42, stratify=y)

```

```

        data_info = f"Завантажено
{len(self.data)} зразків. Тренувальна вибірка:
{len(X_train)}, Тестова: {len(X_test)}."

        model_engine =
ModelEngine(data_processor.preprocessor)

        model_engine.train_model(X_train,
y_train)

        best_params =
model_engine.optimize_model(X_train,
y_train)

        metrics =
model_engine.evaluate_model(X_test, y_test)

        joblib.dump(model_engine.pipeline,
'optimized_pdm_model.pkl')

        decision_support =
DecisionSupport(model_engine.pipeline,
X.columns)

        recommendations = []

        test_df = X_test.copy()

        test_df['actual_class'] = y_test

        indices_to_show = []

        for cls in [0, 1, 2]:

            found = test_df[test_df['actual_class']
== cls].head(3)

            if not found.empty:

                indices_to_show.extend(found.index.tolist())

        if len(indices_to_show) < 3:

            extra = test_df.index[:3 -
len(indices_to_show)].tolist()

            indices_to_show.extend(extra)

        demo_data = X_test.loc[indices_to_show]

```

```

        actual_ruls =
self.data.loc[indices_to_show, 'RUL_days']

        actual_classes =
self.data.loc[indices_to_show, 'target_class']

        for i in range(len(demo_data)):

            sample = demo_data.iloc[[i]]

            rec =
decision_support.generate_recommendation(sa
mple)

            recommendations.append({

                'id': demo_data.index[i],

                'rul_days': actual_ruls.iloc[i],

                'actual_class':
RECOMMENDATION_MAP.get(actual_class
es.iloc[i]),

                'prediction': rec['Prediction'],

                'confidence': rec['Confidence'],

                'reasoning': rec['Reasoning'],

                'sample_data':
sample.to_dict('records')[0]

            })

        return {'data_info': data_info, 'metrics':
metrics, 'best_params': best_params,

                'recommendations':
recommendations }

def run_gui():

    main_system = Main()

    results = main_system.run_analysis()

    root = tk.Tk()

    root.title("СППР: Прогнозна аналітика
обладнання зв'язку")

    root.geometry("1000x800")

```

```

        frame_metrics = tk.Frame(root,
padding="10")

        frame_metrics.pack(fill='x')

        ttk.Label(frame_metrics, text="---
ЗВЕДЕННЯ АНАЛІЗУ ---", font=("Arial", 14,
"bold")).pack(pady=5)

        ttk.Label(frame_metrics,
text=results['data_info'], font=("Arial",
10)).pack(anchor='w')

        metrics_text = "Кінцеві метрики
ОПТИМІЗОВАНОЇ МОДЕЛІ:\n"

        for key, value in results['metrics'].items():

            metrics_text += f" {key}: {value:.4f}\n"

            metrics_text += f" Найкращі параметри:
{results['best_params']}"

            ttk.Label(frame_metrics, text=metrics_text,
justify='left', anchor='w',
foreground='blue').pack(pady=10, fill='x')

            ttk.Label(root, text="---
ПРОГНОЗУВАННЯ ТА ОБҐРУНТУВАННЯ
РІШЕНЬ (ХАІ) ---",

                font=("Arial", 14,
"bold")).pack(pady=10)

            output_text = scrolledtext.ScrolledText(root,
wrap=tk.WORD, width=120, height=35,
font=("Courier New", 10))

            output_text.pack(pady=10, padx=10)

            for rec in results['recommendations']:

                output_text.insert(tk.END,
f"=====
=====\\n")

                output_text.insert(tk.END,
f"ОБЛАДНАННЯ ID (ЗРАЗОК):
{rec['id']}\n", 'header')

                output_text.insert(tk.END, f"-----
-----\\n")

                output_text.insert(tk.END, f"ПРОГНОЗ
РИЗИКУ: {rec['prediction']}\n", 'prediction')

```

```
output_text.insert(tk.END, f"Впевненість
моделі: {rec['confidence']:.2%}\n")
```

```
output_text.insert(tk.END, f"Істинний
RUL: {rec['rul_days']:.1f} днів (Клас:
{rec['actual_class']})\n\n")
```

```
output_text.insert(tk.END,
"ОБҐРУНТУВАННЯ (КЛЮЧОВІ
ФАКТОРИ):\n", 'reasoning_header')
```

```
output_text.insert(tk.END, f"->
{rec['reasoning']}\n")
```

```
output_text.insert(tk.END, "\nСИРІ
ПОКАЗНИКИ:\n", 'data_header')
```

```
for k, v in rec['sample_data'].items():
```

```
    if isinstance(v, (int, float)):
```

```
        output_text.insert(tk.END, f" {k}:
{v:.2f}\n")
```

```
    else:
```

```
        output_text.insert(tk.END, f" {k}:
{v}\n")
```

```
output_text.insert(tk.END,
"=====
=====
=====\\n\\n")
```

```
output_text.tag_config('header',
foreground='black', background='lightgray',
font=("Courier New", 10, "bold"))
```

```
output_text.tag_config('prediction',
foreground='red', font=("Courier New", 12,
"bold"))
```

```
output_text.tag_config('reasoning_header',
foreground='darkgreen', font=("Courier New",
10, "bold"))
```

```
output_text.tag_config('data_header',
foreground='gray', font=("Courier New", 10))
```

```
output_text.config(state=tk.DISABLED) #
Заборонити редагування
```

```
root.mainloop()
```

```
if __name__ == "__main__":
```

```
    run_gui()
```