

УДК: 004.94

*Мудрак П. Р., здобувач вищої освіти,
Бабаков Р. М., професор кафедри інформаційних технологій,
Донецький національний університет імені Василя Стуса*

СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ДІАГНОСТИКИ НЕСПРАВНОСТЕЙ НА ОСНОВІ RANDOM FOREST

Анотація. Розроблено систему автоматичної діагностики несправностей комп'ютерної техніки з використанням алгоритму Random Forest. Реалізовано механізм активного навчання з експертним втручанням. Досягнуто точність моделі 78 % для 20 класів несправностей під час роботи з синтетичним датасетом зі стохастичним шумом.

Ключові слова: машинне навчання, Random Forest, діагностика несправностей, активне навчання, система підтримки прийняття рішень.

Діагностика несправностей комп'ютерної техніки традиційно базується на досвіді майстрів-ремонтників та потребує значних часових витрат. Застосування методів машинного навчання дає змогу автоматизувати цей процес і створити інтелектуальні системи підтримки прийняття рішень (СППР) відповідного профілю. У цій роботі пропонується програмне забезпечення для автоматичної діагностики несправностей на основі симптомів з використанням алгоритму Random Forest та механізму активного навчання.

Розроблена система має багаторівневу програмну архітектуру: рівень представлення (GUI на PyQt6), рівень бізнес-логіки (обробка запитів), рівень машинного навчання (Random Forest), рівень даних (SQLite). Для навчання моделі створено синтетичний датасет обсягом 5 000 записів з додаванням стохастичного шуму: варіативність симптомів ($\pm 30\%$), параметрів ремонту ($\pm 20\%$), додавання помилкових симптомів (20 % ймовірність). Це імітує реальні умови, коли користувач може неточно описати проблему.

Метод Random Forest обрано з таких причин: стійкість до шуму, робота з обмеженими даними (5 000 записів), надання ймовірнісних оцінок для кожного класу. Модель використовує 250 дерев рішень з максимальною глибиною 12. Векторизація ознак виконується відповідно до One-Hot Encoding-симптомів (40 бінарних ознак) та Label Encoding типів пристроїв.

Початкова версія з 50 класами давала точність лише 40–45 % через недостатню кількість прикладів на клас. Після об'єднання схожих несправностей до 20 класів (по 5 на кожен тип пристрою) точність зросла до 78.2 %. Система підтримує 4 типи обладнання: блоки живлення ноутбуків та ПК, струменеві та лазерні принтери.

Реалізовано трипороговий механізм активного навчання: за впевненості не менше 35 % система діагностує несправність автоматично; за впевненості 15–35 % пропонує експерту три найкращі варіанти; за впевненості менше 15 % відкривається діалог експертного введення. Експертні дані зберігаються у трьох варіаціях із різними комбінаціями симптомів для забезпечення мінімальної кількості прикладів кожного класу під час перенавчання [1].

Експериментальне дослідження показало, що 83 % діагностик виконуються автоматично або напівавтоматично (зони ≥ 35 %). Додавання 15 експертних записів (0.3 % від датасету) підвищило точність до 81.8 % (+3.6 %) та покриття автоматичної діагностики до 90 %. Система генерує PDF-звіти з підтримкою української мови через бібліотеку ReportLab та шрифт Arial Unicode [2].

Порівняння з альтернативними методами, що також були досліджені: логістична регресія – 54 %, одиночне дерево рішень – 66 %, нейронна мережа – 61 % (недостатньо даних), Gradient Boosting – 79 % (у 4.7 разів більш повільне навчання). Метод Random Forest забезпечує оптимальний баланс точності, швидкості (2 секунди на навчання, менше 0.1 секунди на прогноз) та інтерпретованості через аналіз важливості ознак [3].

Внаслідок цього розроблено функціональну систему підтримки прийняття рішень для діагностики несправностей комп'ютерної техніки з точністю 78.2 % для 20 класів несправностей. Підтверджено ефективність підходу Human-in-the-Loop Learning для поступового покращення моделі в умовах реального часу. Система може застосовуватися в сервісних центрах, а також для прискорення діагностики та навчання молодих спеціалістів. Перспективи подальшого розвитку розробленої системи: збір реальних даних від сервісних центрів, розширення бази пристроїв, інтеграція з онлайн-сервісами.

Список використаних джерел

1. Settles B. Active Learning Literature Survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009. 67 p.
2. ReportLab User Guide. PDF Generation in Python. URL: <https://www.reportlab.com/docs/reportlab-userguide.pdf> (дата звернення: 16.10.2025).
3. Breiman L. Random Forests. *Machine Learning*. 2001. Vol. 45, № 1. P. 5–32.