

Industrial Transport of Kazakhstan
ISSN 1814-5787 (print)
ISSN 3006-0273 (online)
Vol. 23. Is. 3. Number 91 (2026). Pp. 78–89
Journal homepage: <https://prom.mtgu.edu.kz>
<https://doi.org/10.58420.ptk.2026.91.03.007>
UDC / УДК / ЭОЖ 73.29.41

IIoT-BASED FAULT DETECTION IN A RAIL VEHICLE AIR COMPRESSOR

*S. Sailaubaiuly**

Abai Kazakh National Pedagogical University, Almaty, Kazakhstan.

E-mail: sailaubaiuly@gmail.com

Samat Sailaubaiuly — PhD student, educational programme 8D06102 Information Systems, Abai Kazakh National Pedagogical University, 13 Dostyk Ave., 050010, Almaty, Kazakhstan
E-mail: sailaubaiuly@gmail.com, <https://orcid.org/0009-0002-0458-3242>.

© S. Sailaubaiuly

Abstract. Scheduled maintenance does not always reflect the actual condition of equipment, and the built-in threshold signals fire only after a failure has begun. We assess, at event level, whether anomaly detection models applied to IIoT telemetry detect a failure earlier than the built-in threshold method or with fewer false alarms. The data are the open MetroPT-3 dataset from the air production unit of a metro train (February–August 2020, 1,516,948 records, 15 signals, CC BY 4.0 licence, 4 failures from company reports) and the MetroPT2 dataset (2022, 2 failures) for a transfer test. The methods comprised 38 features in 10-minute windows, a strictly chronological split, Isolation Forest, One-Class SVM, a fully connected autoencoder, an EWMA control chart and the built-in LPS rule; thresholds were chosen on validation data for a target number of false alarms per day, and each model was run with 5 random seeds. At the pre-set operating point of 0.1 false alarms per day the models detected on average 0–0.6 of 4 failures, EWMA 1, and no method alarmed before the reported start. With thresholds calibrated to 0.5 false alarms per day, the models and EWMA detected on average 3.8–4.0 failures at 0.15–0.19 false alarms per day, the built-in rule 3 failures at 0.66; median lead times were negative. The models ranked windows near failures better than EWMA (PR-AUC 0.40–0.45 versus 0.24). Models trained on 2020 data kept this ability on 2022 data (PR-AUC 0.69–0.73), but the 2020 threshold was too high. The novelty is an event-level comparison with the built-in rule on real metro compressor data, with a chronological protocol and a transfer test. In practice, the threshold can be set from the false-alarm rate a maintenance team accepts. Early warning was not demonstrated; thresholds must be recalibrated for each asset.

Keywords: IIoT, anomaly detection, condition-based maintenance, air compressor, event-level evaluation, false alarms

For citation: Sailaubaiuly S. (2026). IIoT-Based Fault Detection in a Rail Vehicle Air Compressor // Industrial Transport of Kazakhstan. Vol. 23. No. 91. Pp. 78–89. <https://doi.org/10.58420.ptk.2026.91.03.007> (In Kaz.).

Conflict of interest: The author declares no conflict of interest.

IIoT ДЕРЕКТЕРІ БОЙЫНША РЕЛЬСТІК КӨЛІК КОМПРЕССОРЫНЫҢ АҚАУЛАРЫН АНЫҚТАУ

*С. Сайлаубайұлы**

Абай атындағы Қазақ ұлттық педагогикалық университеті, Алматы, Қазақстан.

E-mail: sailaubaiuly@gmail.com

Сайлаубайұлы Самат — PhD докторанты, 8D06102 – «Ақпараттық жүйелер» білім беру бағдарламасы, Абай атындағы Қазақ ұлттық педагогикалық университеті, Достық даңғылы, 13, 050010, Алматы, Қазақстан.

E-mail: sailaubaiuly@gmail.com, <https://orcid.org/0009-0002-0458-3242>

© Сайлаубайұлы С.

Аннотация. Жоспарлы техникалық қызмет көрсету жабдықтың нақты күйін әрдайым ескермейді, ал штаттық шекті сигналдар ақау басталғаннан кейін іске қосылады. Зерттеудің мақсаты – ПоТ телеметриясы бойынша аномалияларды анықтау модельдері ақауды штаттық шекті әдіске қарағанда ертерек немесе азырақ жалған дабылмен анықтай ала ма – осыны оқиғалар деңгейінде бағалау. Деректер – метро пойызының ауа өндіру агрегатының ашық MetroPT-3 жиыны (2020 ж. ақпан–тамыз, 1 516 948 жазба, 15 сигнал, CC BY 4.0 лицензиясы, кәсіпорын есептерінде тіркелген 4 ақау) және модельді басқа кезеңге ауыстыруды тексеруге арналған MetroPT2 жиыны (2022 ж., 2 ақау). Әдістер: 10 минуттық терезелердегі 38 белгі, қатан хронологиялық бөлу, Isolation Forest, One-Class SVM, толық байланысты автокодтаушы, EWMA бақылау картасы және LPS штаттық ережесі; шекті мән жалған дабылдардың берілген тәуліктік деңгейі бойынша валидацияда тандалды; әр модель бес түрлі кездейсоқ бастапқы мәнмен (random seed) іске қосылды. Нәтижелер: алдын ала белгіленген тәулігіне 0,1 жалған дабыл деңгейінде модельдер 4 ақаудың орта есеппен 0–0,6-сын, EWMA 1-ін анықтады және ешбір әдіс ақауды есептегі басталу уақытынан бұрын анықтаған жоқ. Тәулігіне 0,5 деңгейде модельдер мен EWMA орта есеппен 3,8–4,0 ақауды тәулігіне 0,15–0,19 жалған дабылмен анықтады, ал штаттық ереже 3 ақауды тәулігіне 0,66 жалған дабылмен анықтады; ескерту уақытының медианасы барлық әдістерде теріс болды. Модельдер ақау маңындағы терезелерді EWMA-ға қарағанда жақсырақ ажыратты (PR-AUC 0,40–0,45, EWMA үшін 0,24). 2020 жылғы модельдер 2022 жылғы деректерде ажырату қабілетін сақтады (PR-AUC 0,69–0,73), бірақ 2020 жылғы шекті мән тым жоғары болды. Ғылыми жаңалығы – метро компрессорының нақты пайдалану деректерінде модельдерді штаттық ережемен оқиғалар деңгейінде, хронологиялық хаттамамен және кезеңдер арасында ауыстыруды тексерумен салыстыру. Практикалық маңызы – дабыл шегін техникалық қызмет көрсету қызметі қабылдай алатын жалған дабылдар саны бойынша таңдау әдістемесі. Қорытынды: ерте ескерту байқалған жоқ, ал шекті мәнді нақты нысанның деректерінде қайта калибрлеу қажет.

Түйін сөздер: ПоТ, аномалияларды анықтау, күйі бойынша техникалық қызмет көрсету, компрессорлық агрегат, оқиғалар деңгейіндегі бағалау, жалған дабылдар

Дәйексөз үшін: Сайлаубайұлы С. (2026). ПоТ деректері бойынша рельстік көлік компрессорының ақауларын анықтау // Қазақстан өндіріс көлігі. Том. 23. № 91. 78–89 бет. <https://doi.org/10.58420.ptk.2026.91.03.007> (Қаз. тіл.).

Мүдделер қақтығысы: Автор мүдделер қақтығысының жоқтығын мәлімдейді.

ОБНАРУЖЕНИЕ ОТКАЗОВ КОМПРЕССОРА РЕЛЬСОВОГО ТРАНСПОРТА ПО ДАННЫМ ПоТ

*С. Сайлаубайұлы**

Казахский национальный педагогический университет имени Абая, Алматы, Казахстан.

Сайлаубайұлы Самат — PhD-докторант, образовательная программа 8D06102 – «Информационные системы», Казахский национальный педагогический университет имени Абая, пр. Достык, 13, 050010, Алматы, Казахстан

E-mail: sailaubaiuly@gmail.com, <https://orcid.org/0009-0002-0458-3242>.

© Сайлаубайұлы С.



Аннотация. Плановое техническое обслуживание не всегда учитывает фактическое состояние оборудования, а штатные пороговые сигналы срабатывают уже после начала отказа. Цель исследования – оценить на уровне событий, позволяют ли модели обнаружения аномалий по телеметрии ПоТ выявлять отказ раньше штатного порогового метода или с меньшим числом ложных тревог. Данные – открытый набор MetroPT-3 по воздухоприготовительному агрегату (APU) поезда метро (февраль–август 2020 г., 1 516 948 записей, 15 сигналов, лицензия CC BY 4.0, 4 отказа по отчётам предприятия) и набор MetroPT2 (2022 г., 2 отказа) для проверки переносимости. Методы: 38 признаков в 10-минутных окнах, строго хронологическое разбиение, Isolation Forest, One-Class SVM, полносвязный автокодировщик, контрольная карта EWMA и штатное правило LPS; порог выбирался на валидации по заданному числу ложных тревог в сутки; каждая модель запускалась с пятью значениями random seed. Результаты: при заранее заданной рабочей точке 0,1 ложной тревоги в сутки модели обнаружили в среднем 0–0,6 из 4 отказов, EWMA – 1, и ни один метод не дал тревоги раньше начала отказа по отчёту. При калибровке на 0,5 ложной тревоги в сутки модели и EWMA обнаружили в среднем 3,8–4,0 из 4 отказов при 0,15–0,19 ложной тревоги в сутки, а штатное правило – 3 отказа при 0,66 ложной тревоги в сутки; медиана упреждения у всех методов отрицательна. Модели лучше EWMA ранжировали окна вблизи отказов (PR-AUC 0,40–0,45 против 0,24). Модели, обученные на данных 2020 г., сохранили ранжирующую способность на данных 2022 г. (PR-AUC 0,69–0,73), но порог 2020 г. оказался слишком высоким. Научная новизна – событийное сравнение моделей со штатным правилом на реальных эксплуатационных данных компрессора метро при хронологическом протоколе и проверке переноса между периодами. Практическое значение – методика выбора порога тревоги по допустимому для службы обслуживания числу ложных тревог. Вывод: раннее предупреждение не подтверждено, а порог необходимо калибровать на данных конкретного объекта.

Ключевые слова: ПоТ, обнаружение аномалий, обслуживание по состоянию, компрессорный агрегат, событийная оценка, ложные тревоги

Для цитирования: Сайлаубайулы С. (2026). Обнаружение отказов компрессора рельсового транспорта по данным ПоТ // Помышленный транспорт Казахстана. Т. 23. No. 91. Стр. 78–89. <https://doi.org/10.58420.ptk.2026.91.03.007> (На каз.).

Конфликт интересов: Автор заявляет об отсутствии конфликта интересов.

Кіріспе.

Өнеркәсіптік және қалалық рельстік көліктің жылжымалы құрамында пневматикалық жүйелер тежеу мен көмекші жетектерді сығылған ауамен қамтамасыз етеді, ал ауа өндіру агрегатының (компрессордың) істен шығуы пойыздың желіден алынуына әкелуі мүмкін. Мұндай жабдыққа техникалық қызмет көрсету көбіне жоспарлы кесте бойынша жүргізіледі және агрегаттың нақты күйін ескермейді, сондықтан ақау қызмет көрсету аралығында пайда болуы мүмкін. Күйі бойынша қызмет көрсету мен болжамды қызмет көрсетудің мақсаты – бұл олқылықты датчик деректері арқылы жою (Zio, 2022: 108119; Carvalho және т.б., 2019: 106024).

Қазіргі жылжымалы құрамда датчиктер деректерін борттық ПоТ құрылғылары жинайды, бірақ олармен қатар штаттық шекті сигналдар да жұмыс істейді: мысалы, қысым белгіленген деңгейден төмен түскенде іске қосылатын сигнал. Мұндай сигнал ақау басталғаннан кейін ғана дабыл береді. Машиналық оқыту модельдерінен көбіне ерте анықтау күтіледі (Zio, 2022: 108119), алайда қаралған жұмыстарда бұл күту штаттық ережемен салыстырылып, жалған дабылдар құнымен бірге тексерілмеген (төменде қараңыз).

Мұндай салыстырудың нәтижесі бағалау әдістемесіне қатты тәуелді. Уақыттық қатарлардағы аномалияларды анықтау бойынша жарияланған көптеген нәтижелер нүктелік түзету сияқты тәсілдердің есебінен асыра бағаланатыны көрсетілген (Kim және т.б., 2022:

7194–7201), ал кең тараған эталондық жиындардың өзі сынға ұшыраған (Wu, Keogh, 2021). Техникалық қызмет көрсету қызметі үшін маңыздысы – терезелер деңгейіндегі метрикалар емес, нақты ақаудың қаншалықты ерте анықталатыны және тәулігіне қанша жалған дабыл берілетіні. Сондықтан модельдерді оқиғалар деңгейінде, шекті мәнді тест кезеңіне дейін бекітіп, уақыт бойынша бөлінген деректерде бағалау қажет.

Ашық MetroPT деректер жиындары мұндай бағалауды қайталанатын түрде жүргізуге мүмкіндік береді: оларда метро пойызының ауа өндіру агрегатының пайдалану кезеңіндегі телеметриясы және кәсіпорын есептерінде тіркелген ақаулар бар (Veloso және т.б., 2022a). Бұл деректер Қазақстанның өнеркәсіптік көлігіне тікелей қатысты емес, бірақ ұсынылған бағалау хаттамасы жабдық түріне тәуелсіз, сондықтан оны кейін өнеркәсіптік локомотивтер мен вагондардың пневматикалық жүйелерінің деректерінде қолдануға болады; мұндай көшірудің нәтижесі бөлек тексеруді қажет етеді.

Зерттеу нысаны – метро пойызы ауа өндіру агрегатының IoT телеметриясы. Зерттеу пәні – уақыттық қатарлардағы аномалияларды анықтау әдістері және дабыл шегін таңдау ережесі. Мақсат – аномалияларды анықтау модельдері ақауды штаттық шекті әдіске қарағанда ертерек немесе азырақ жалған дабылмен анықтай ала ма – осыны оқиғалар деңгейінде бағалау. Міндеттер: ашық деректерге аудит жүргізу; хронологиялық бағалау хаттамасын құру; модельдерді штаттық ережемен және статистикалық бақылау картасымен салыстыру; модельдердің 2020 жылдан 2022 жылға ауыстырылуын тексеру. Мақалада агрегаттың қалдық ресурсын (RUL) болжау қарастырылмайды: қолжетімді деректерде агрегаттың «істен шыққанға дейінгі» толық циклдері жоқ.

Жұмыста екі болжам тексеріледі. Біріншісі – кем дегенде бір модель штаттық ережеге қарағанда ақауды ертерек ескертеді және бұл ретте тәулігіне жалған дабылдар саны одан көп болмайды. Екіншісі – 2020 жылғы деректерде оқытылған модель 2022 жылғы ақауларды анықтау қабілетін сақтайды. Бірінші болжам ерте ескерту туралы кең тараған күтуді тексереді, ал екіншісі модельді жабдықтың басқа пайдалану кезеңіне қайта оқытусыз қолдануға бола ма деген практикалық сұраққа жауап береді.

Мақала келесі түрде құрылған: әдебиетке шолу бөлімінде MetroPT деректері бойынша бұрынғы жұмыстар мен бағалау әдістемесіне қатысты сын талданып, зерттеу сұрақтары тұжырымдалады; «Материалдар мен әдістер» бөлімінде деректер аудиті, белгілер, модельдер, шекті мәнді калибрлеу және оқиғалар деңгейіндегі көрсеткіштер сипатталады; «Нәтижелер» бөлімінде 2020 жылғы тест кезеңі мен 2022 жылғы деректердегі нәтижелер келтіріледі; «Талқылау» және «Қорытынды» бөлімдерінде практикалық салдары мен шектеулері қарастырылады.

MetroPT деректері Порту қаласының метро операторымен бірлескен жобада жинақталған. Veloso және әріптестері деректер жиынының құрылымын, сигналдарын және ақаулар туралы есептерді сипаттады (Veloso және т.б., 2022a). Осы агрегаттың 2020 жылғы деректерін Davari және әріптестері терең оқытуға негізделген аномалияларды анықтау үшін пайдаланды (Davari және т.б., 2021: 1–10), кейін аномалияларды анықтау мен түсіндіру тәсілі ұсынылды (Davari және т.б., 2024: 358–364). Агрегаттың ақауын пайдалану жағдайында анықтау бойынша ертерек жұмыс та бар (Bartos және т.б., 2020: 61–74).

Кейінгі жұмыстар ағындық оқытуға бағытталған. Meira және әріптестері теміржол жүйелері үшін ағындық деректерге негізделген болжамды қызмет көрсету шеңберін ұсынды (Meira және т.б., 2023: 1087–1102); Zafra және әріптестері MetroPT-3 деректерінде болжау, автокодтаушы және бірнеше сигналдың бір мезгілде ауытқуын талап ететін ережені біріктірді (Zafra және т.б., 2024: 339–351); García-Méndez және әріптестері 2022 жылғы деректерде түсіндірмелі инкременттік жіктеуіштерді қолданды (García-Méndez және т.б., 2025). Соңғы жұмыста ақаудың алдындағы 2 сағат ақау ретінде белгіленеді және тек ақаулар маңындағы аралықтар сақталады, гиперпараметрлер кездейсоқ ішкі іріктемеде таңдалады; нәтижелер нүктелік F-өлшеммен беріледі.

Уақыттық қатарлардағы аномалияларды анықтау әдістерін бағалау әдістемесінің өзі де сынға ұшырап отыр. Kim және әріптестері нүктелік түзету (point adjustment) нәтижелерді асыра бағалайтынын көрсетті (Kim және т.б., 2022: 7194–7201), ал Wu мен Keogh кең таралған эталондық жиындардың кемшіліктері прогресс елесін тудыратынын атап өтті (Wu, Keogh, 2021). Hundman және әріптестері ғарыш аппараттары телеметриясы үшін оқиғалар деңгейіндегі бағалау мен динамикалық шекті мәнді қолданды (Hundman және т.б., 2018: 387–395). Теңгерімсіз деректерде дәлдік–толықтық қисығы ROC қисығына қарағанда көбірек ақпарат береді (Davis, Goadrich, 2006: 233–240; Saito, Rehmsmeier, 2015).

Қаралған әдебиетте MetroPT деректері бойынша (1) модельдерді агрегаттың штаттық шекті ережесімен салыстыру, (2) ақауға дейінгі ескерту уақыты мен тәулігіне жалған дабылдар саны сияқты оқиғалар деңгейіндегі көрсеткіштер, (3) қатаң хронологиялық хаттама және (4) модельді бір пайдалану кезеңінен екіншісіне ауыстыру бір жұмыста кездеспеді. Осыған байланысты зерттеу сұрағы: IoT телеметриясы бойынша аномалияларды анықтау модельдері компрессор ақауын штаттық шекті әдіске қарағанда ертерек немесе азырақ жалған дабылмен анықтай ала ма және бұл артықшылық басқа пайдалану кезеңінде сақтала ма?

Материалдар мен әдістер.

Зерттеу бес кезеңнен тұрады: деректер аудиті; белгілерді құру; модельдерді оқыту және шекті мәнді калибрлеу; оқиғалар деңгейінде бағалау; модельді басқа кезеңге ауыстыруды тексеру. Екі болжам тексерілді. H1: кем дегенде бір модель штаттық ережеге қарағанда ертерек ескертеді және тәулігіне жалған дабылдар саны одан көп емес. H2: 2020 жылғы деректерде оқытылған модель 2022 жылғы ақауларды анықтауды сақтайды.

Деректер. Негізгі жиын – MetroPT-3 (Davari және т.б., 2023; CC BY 4.0 лицензиясы). Аудит көрсеткендей, CSV файлы 1 516 948 жазбадан тұрады және 10 с қадаммен жазылған (0,1 Гц), ал сипаттамада 1 Гц жиілігі және 15 169 480 нүкте көрсетілген; жазбалар саны сипаттамадағы нүктелер санының оннан біріне тең. Өткізіп алынған мәндер жоқ, бірақ 60 с-тан ұзақ 331 үзіліс бар, олардың жалпы ұзақтығы 909,5 сағ; ең ұзағы – 2020 жылғы 25–27 сәуірдегі 48 сағаттық үзіліс. Деректер 2020 жылғы 1 ақпаннан 1 қыркүйекке дейінгі кезеңді қамтиды. Жеті аналогтық сигнал (TP2, TP3, H1, DV_pressure, Reservoirs, Oil_temperature, Motor_current) және сегіз дискретті сигнал (COMP, DV_electric, Towers, MPG, LPS, Pressure_switch, Oil_level, Caudal_impulses) бар; файлдағы баған атаулары осылай жазылған. Ақаулар туралы кәсіпорын есептері деректер бетінде берілген: 4 ақау, барлығы «ауаның ағуы» түрінде (1-кесте).

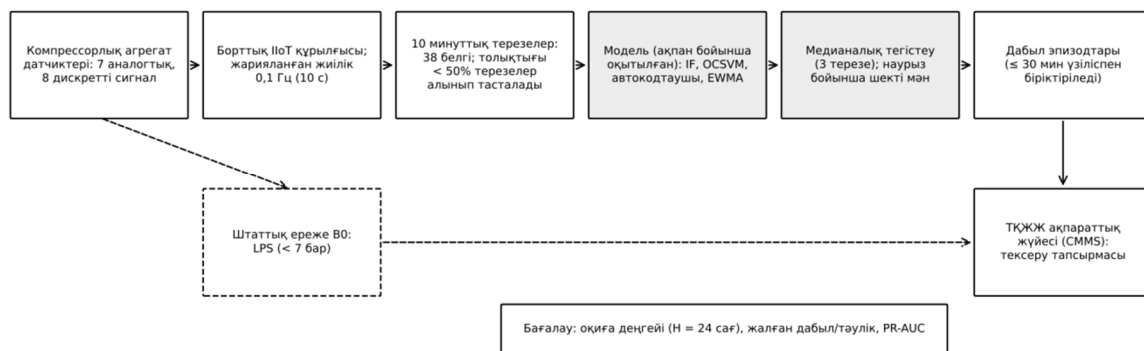
Кесте 1 — Деректер жиыны, бөлу және ақаулар

Көрсеткіш	Мәні	Ескертпе
Кезең; жазбалар саны	01.02–01.09.2020; 1 516 948	қадамы 10 с (0,1 Гц)
Үзілістер > 60 с; жалпы ұзақтығы	331; 909,5 сағ	ең ұзағы 48 сағ
Оқыту (ақпан) / валидация (наурыз) / тест (сәуір–тамыз), терезе	3552 / 3809 / 17895	24,7 / 26,5 / 124,3 тәулік
Ақау 1	18.04.2020 00:00 – 23:59	ауаның ағуы
Ақау 2	29.05.2020 23:30 – 30.05.2020 06:00	ауаның ағуы; есепте қызмет көрсету күні «30 Apr» деп кәте берілген
Ақау 3	05.06.2020 10:00 – 07.06.2020 14:30	ауаның ағуы; деректермен қамтылуы 92%
Ақау 4	15.07.2020 14:30 – 19:00	ауаның ағуы
MetroPT2 (ауыстыруды тексеру)	28.04–28.07.2022; 7 116 940 жазба, 1 Гц	0,1 Гц-ке келтірілді; 2 ақау (ауа, май ағуы)

Дереккөз: MetroPT-3 және MetroPT2 деректері мен беттері негізінде автордың аудиті. Терезе – 10 минуттық аралық; тәулік – деректермен қамтылған уақыт.

Сигналдар аяқталу уақытымен белгіленетін 10 минуттық қабаттаспайтын терезелерге топтастырылды; әр терезе тек өз ішіндегі мәндерді пайдаланады, сондықтан болашақтан ақпарат кірмейді. Әр аналогтық сигнал үшін орташа мән, стандартты ауытқу, минимум және максимум, әр дискретті сигнал үшін белсенді күйдің үлесі есептелді. Сонымен қатар қозғалтқыш тогы 5,5 А-дан жоғары болған уақыт үлесі (жүктемедегі жұмыс үлесі; сипаттамада бос жүріс шамамен 4 А, жүктеме шамамен 7 А деп берілген) және СОМР сигналының ауысу саны қосылды. Барлығы 38 белгі. Өлшеулердің 50 %-ынан азы бар терезелер алынып тасталды.

Бөлу және деректердің ағып кетуін (data leakage) болдырмау. Бөлу қатаң хронологиялық: оқыту – 2020 жылғы ақпан (деректер бетінде ұсынылғандай, бірінші ай), валидация – наурыз (ақаулар жоқ, тек шекті мәнді таңдау үшін), тест – сәуір–тамыз. Масштабтау параметрлері (медиана мен кuartиль аралығы; стандартталған мәндер ± 50 -пен шектеледі) бағаланып, модельдер тек оқыту іріктемесінде оқытылды. В0-ден басқа барлық әдістерге себептік өңдеу қолданылды: соңғы 3 терезе бойынша медианалық тегістеу; 30 минуттан аз үзілісі бар дабылдарды бір эпизодқа біріктіру барлық әдістерге ортақ (1-сурет).



Сур. 1. IoT деректерін өңдеу және дабыл қалыптастыру схемасы

Дереккөз: автор құрастырған. IF – Isolation Forest; OCSVM – One-Class SVM; EWMA – экспоненциалды өлшенген жылжымалы орташа; ТҚЖЖ – техникалық қызмет көрсету және жөндеу; CMMS – техникалық қызмет көрсетуді басқару жүйесі.

Модельдер мен салыстыру негіздері. В0 – агрегаттың штаттық ережесі: терезеде LPS сигналы (қысым 7 бардан төмен) белсенді болса, дабыл беріледі; бұл ереже калибрленбейді. В1 – жүктемедегі жұмыс үлесі бойынша бір жақты EWMA бақылау картасы ($\alpha = 0,1$), мәндері оқыту кезеңінің орташа мәні мен стандартты ауытқуы бойынша стандартталады (Roberts, 1959: 239–250); $L \cdot \sigma$ бақылау шектерінің орнына оның шекті мәні модельдердікі сияқты калибрленеді. Модельдер: Isolation Forest (200 ағаш; аномалия бағасы – score_samples мәнінің теріс таңбалысы) (Liu және т.б., 2008: 413–422); One-Class SVM ($\nu = 0,01$, RBF ядросы, $\gamma = \text{scale}$, оқыту іріктемесінен 3000 терезелік ішкі іріктеме; баға – decision_function мәнінің теріс таңбалысы) (Schölkopf және т.б., 2001: 1443–1471); 16–4–16 құрылымды толық байланысты автокодтаушы (ReLU, Adam, ең көбі 400 итерация, ерте тоқтату; баға – қалпына келтірудің орташа квадраттық қатесі). Әр модель 0–4 кездейсоқ бастапқы мәндермен (random seed) бес рет оқытылды.

Шекті мәнді калибрлеу. Шекті мән валидация кезеңінде таңдалады: тегістелген бағалардың 90–100 % квантильдері ішінен валидациядағы дабыл эпизодтарының тәуліктік саны берілген деңгейден (0,05; 0,1; 0,2; 0,5 тәулігіне) аспайтын ең кіші мән алынады. Тест кезеңінде шекті мән өзгертілмейді. Осылайша, жалған дабылдардың рұқсат етілген саны инженерлік талап ретінде алдын ала белгіленеді, ал оның нақты деңгейін техникалық қызмет көрсету қызметі анықтауы тиіс.

Көрсеткіштер. Егер [басталуы – Н; аяқталуы] аралығында кемінде бір дабыл болса, ақау анықталды деп саналады; негізгі көкжиек $H = 24$ сағ, сезімталдық үшін 6, 12 және 48 сағ қарастырылды. Ескерту уақыты – есептегі басталу уақыты мен алғашқы дабыл арасындағы айырма (оң мән – ақаудан бұрын). Жалған дабыл – [басталуы – Н; аяқталуы + 24 сағ] аймақтарынан тыс басталған эпизод; олардың саны осы аймақтардан тыс деректермен қамтылған уақыттың тәулігіне келтіріледі. Терезелер деңгейінде дәлдік–толықтық қисығы бойынша орташа дәлдік (average precision, мақалада PR-AUC) есептелді: оң класс – [басталуы – Н; аяқталуы] аралығындағы терезелер, ақаудан кейінгі 24 сағаттық терезелер есептен алынып тасталды; оң кластың үлесі тестте 0,056, бұл кездейсоқ PR-AUC деңгейі. Есептеу кідірісі белгілерді есептемей, тест іріктемесін бағалау уақытын терезелер санына бөлу арқылы өлшенді.

Басқа кезеңге ауыстыруды тексеру (H2). MetroPT2 жиыны (Veloso және т.б., 2022b; CC BY 4.0; 2022 жылғы 28 сәуір – 28 шілде, 1 Гц; MD5 бақылау сомасы көрсетілгенмен сәйкес келді) әр 10 секундтық аралықтағы бірінші өлшемді алу арқылы 0,1 Гц-ке келтірілді. Екі баптау салыстырылды: (а) 2020 жылғы модельдер мен шекті мәндер өзгеріссіз қолданылды; (б) модельдер сол хаттама бойынша 2022 жылғы деректерде қайта оқытылды (оқыту 28.04–15.05, валидация 15.05–01.06). Екі жағдайда да тест кезеңі бірдей – 2022 жылғы 1 маусым – 28 шілде, 2 ақау: ауаның ағуы (04.06) және майдың ағуы (11.07–14.07).

Қайталанушылық. Есептеулер Python 3.13.16, numpy 2.5.3, pandas 3.0.5, scikit-learn 1.9.1 кітапханаларымен Intel Xeon 2,80 ГГц процессорында (2 виртуалды ядро, 7 ГБ жад, GPU жоқ) орындалды; белгілер есептелгеннен кейін негізгі эксперимент 28 с алды. Код, конфигурация, деректер мұрағатының SHA-256 бақылау сомасы, MetroPT2 файлының MD5 бақылау сомасы және барлық іске қосулардың нәтижелері сақталды; кодты құрылымдық өзгертуден кейін қайта жүргізгенде 580 нәтиже жазбасының бірде-бір мәні өзгермеді.

Нәтижелер.

Негізгі нәтижелер 2-кестеде келтірілген. Алдын ала негізгі деп белгіленген тәулігіне 0,1 жалған дабыл деңгейінде модельдер 4 ақаудың орта есеппен тек 0–0,6-сын анықтады: One-Class SVM бірде-бірін, Isolation Forest 0,4-ін, автокодтаушы 0,6-сын, ал EWMA 1-ін; 0,1 және 0,2 деңгейлерінде ешбір әдістің бірде-бір іске қосылуында дабыл ақаудың есептегі басталуынан бұрын берілмеді. Деңгей тәулігіне 0,5-ке дейін жұмсартылғанда модельдер мен EWMA орта есеппен 3,8–4,0 ақауды анықтады, ал тесттегі нақты жалған дабылдар деңгейі тәулігіне 0,15–0,19 болды. Штаттық ереже 3 ақауды анықтады, бірақ 114,5 тәулік ішінде 75 жалған дабыл эпизоды берді (тәулігіне 0,655); бұл көрсеткіш жеке тексеру скриптімен қайта есептеліп, сәйкес келді.

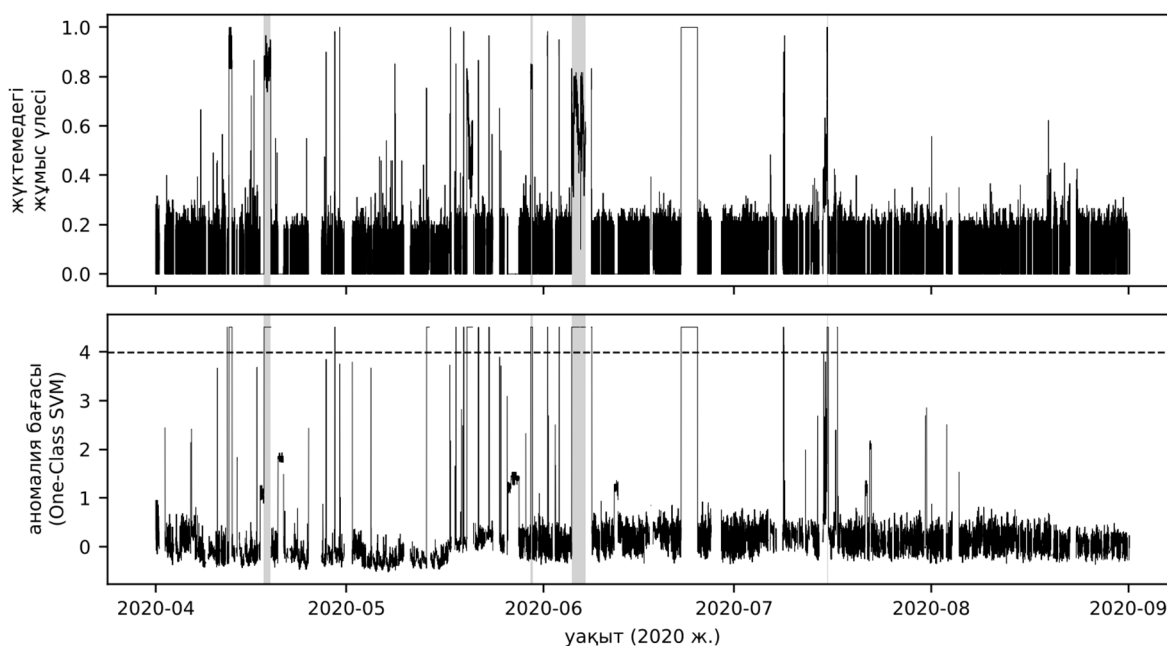
Кесте 2 — Ақауларды анықтау нәтижелері, MetroPT-3, $H = 24$ сағ

Әдіс	Ақау (4-тен), 0,1/тәул.	Ақау (4-тен), 0,5/тәул.	Ескерту медианасы, сағ (0,5)	Жалған дабыл/тәул. (0,5)	PR-AUC
B0: штаттық LPS ережесі	3 (калибренбейді)	3	-3,00	0,655	0,059
B1: EWMA	1,0	4,0	-0,33	0,183	0,235
Isolation Forest	$0,4 \pm 0,5$	$3,8 \pm 0,4$	-0,25	$0,150 \pm 0,018$	$0,400 \pm 0,036$
One-Class SVM	0,0	4,0	-0,25	$0,169 \pm 0,004$	$0,446 \pm 0,005$
Автокодтаушы	$0,6 \pm 0,8$	$3,8 \pm 0,4$	-0,17	$0,194 \pm 0,052$	$0,432 \pm 0,039$

Дереккөз: автордың ашық деректерді талдау нәтижелері. Мәндер – 5 іске қосу бойынша орташа $\pm$ стандартты ауытқу (ауытқу нөлге тең болса көрсетілмейді); 0,1 және 0,5 – валидацияда берілген тәуліктік жалған дабыл деңгейі. Ескерту медианасы – іске қосулардағы медианалардың медианасы; теріс мән дабыл ақау басталғаннан кейін берілгенін білдіреді.

Ескерту уақытының медианасы барлық әдістерде теріс: тәулігіне 0,5 деңгейде дабыл ақау басталғаннан кейін 10–20 минутта беріледі, бұл терезенің аяқталу уақытымен белгіленуі мен үш терезелік тегістеудің құрылымдық кешігуімен шамалас. Жекелеген он мәндер (0,8–13,7 сағ) тек 4-ақауға қатысты: автокодтаушының 4 іске қосылуында, One-Class SVM-нің 2 іске қосылуында және EWMA-да. Штаттық ереже 1-ақауды басталуынан 19,2 сағ бұрын белгіледі, бірақ тәулігіне 0,66 жалған дабыл жиілігінде кез келген 24 сағаттық терезеде кемінде бір дабылдың болу ықтималдығы шамамен 0,48 ($1 - e^{(-0,655)}$), сондықтан бұл сәйкестік кездейсоқ болуы мүмкін. Н1 болжамы (ертерек және жалған дабылдар көп емес) расталмады. Жалған дабылдар бойынша артықшылық жеке нәтиже ретінде байқалды: тәулігіне 0,15–0,19 жалған дабылмен 3,8–4 ақау, штаттық ережеде 0,66 жалған дабылмен 3 ақау. Тәулігіне 0,1 және 0,2 деңгейлерде EWMA модельдерден де жақсы болды (1 және 4 ақау, модельдерде 0–0,6 және 1,4–3,0).

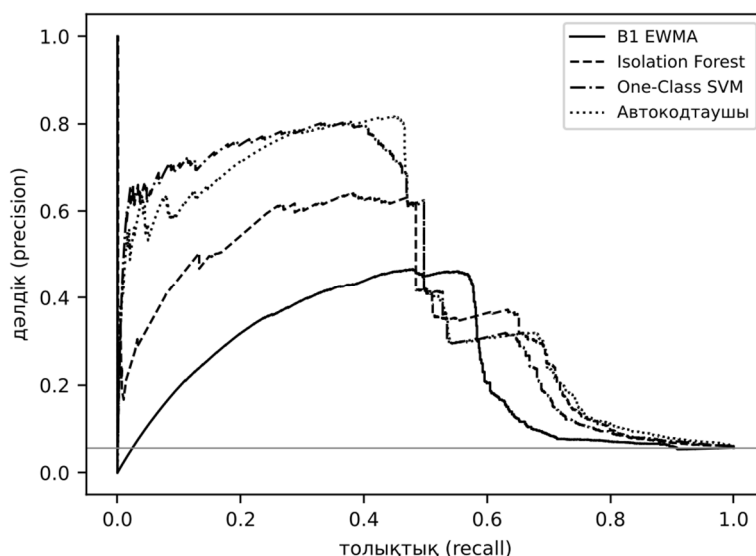
Терезелер деңгейінде модельдер ақаулар маңындағы аралықтарды EWMA-ға қарағанда әлдеқайда жақсы ажыратты (PR-AUC 0,40–0,45, EWMA үшін 0,24; кездейсоқ деңгей 0,056; 2-сурет, 3-сурет). Абляция көрсеткендей, дискретті сигналдарды алып тастау қатаң шекті мәнде анықтауды арттырады (Isolation Forest: 1,4 және 0,4 ақау; автокодтаушы: 1,0 және 0,6), бірақ PR-AUC мәнін төмендетеді (Isolation Forest: 0,382 және 0,400). Белгілерді есептемегенде бір терезені бағалау уақыты 0,001–0,008 мс, яғни бағалаудың өзі шектеуші фактор емес.



Сур.2. Жүктемедегі жұмыс үлесі және One-Class SVM аномалия бағасы, тест кезеңі 2020 ж.

Дереккөз: автордың есептеулері (random seed 0). Сұр жолақтар – есептердегі ақау аралықтары; үзік сызық – тәулігіне 0,5 жалған дабыл деңгейіне калибрленген шекті мән; баға мәндері шекті мәнің үш еселенген деңгейінде қиылған.

2-суретте 2020 жылғы 22–25 маусымда шамамен 62 сағат бойы агрегаттың үздіксіз жүктемеде жұмыс істегені көрінеді; бұл аралық ақаулар туралы есептерде жоқ, бірақ One-Class SVM, автокодтаушы және EWMA оны аномалия ретінде белгілейді (Isolation Forest, seed 0 – белгілемейді). Мұндай эпизодтар есептеуде жалған дабыл болып саналады, дегенмен олар есепке алынбаған нақты ауытқулар болуы мүмкін; бұл ақау таңбаларының (labels) сапасына қатысты шектеу.



Сур.3. Тест кезеңіндегі дәлдік–толықтық қисықтары, $N = 24$ сағ

Дереккөз: автордың есептеулері (random seed 0). Көлденең сұр сызық – кездейсоқ деңгей (0,056).

Басқа кезеңге ауыстыру нәтижелері 3-кестеде берілген. 2020 жылғы модельдер 2022 жылғы деректерде терезелерді жақсы ажыратты (PR-AUC 0,69–0,73; 2022 жылғы кездейсоқ деңгей 0,063; 2022 жылғы деректерде оқытылған модельдерде 0,64–0,74), бірақ 2020 жылғы шекті мән тым жоғары болды: тәулігіне 0,1 деңгейде бірде-бір ақау анықталмады, ал 0,5 деңгейде ақаулар басталғаннан 15–21 сағ кейін ғана анықталды. H2 ішінара расталды: ажырату қабілеті сақталды, бекітілген шекті мәнмен анықтау сақталмады. 2022 жылғы деректерде қайта оқыту екі ақауды да анықтауға мүмкіндік берді; One-Class SVM 0,1 деңгейде екі ақауды да басталуынан 1,7 сағат бұрын анықтады (тәулігіне 0,13 жалған дабыл). Бұл нәтиже тек екі оқиғаға негізделгендіктен, оны дәлел ретінде қарастыруға болмайды. Штаттық ереже 2022 жылы 2020 жылмен салыстырғанда тәулігіне шамамен 10 есе аз жалған дабыл берді (0,066 және 0,655), бұл кезеңдер арасында пайдалану режимі өзгеруі мүмкін екенін меңзейді.

Кесте 3 — 2022 жылғы деректерде басқа кезеңге ауыстыруды тексеру, $N = 24$ сағ

Баптау	Әдіс	Ақау (2-ден), 0,1 / 0,5	Жалған дабыл/тәул. (0,5)	PR-AUC
Ереже	B0: LPS	2	0,066	0,116
2020 → 2022	EWMA	0,0 / 1,0	0,000	0,247
2020 → 2022	Isolation Forest	0,0 / 1,2 ± 0,7	0,000	0,707 ± 0,003
2020 → 2022	One-Class SVM	0,0 / 2,0	0,000	0,734 ± 0,001
2020 → 2022	Автокодтаушы	0,0 / 1,6 ± 0,8	0,000	0,689 ± 0,048
2022 деректерінде оқыту	EWMA	1,0 / 2,0	0,154	0,247
2022 деректерінде оқыту	Isolation Forest	2,0 / 2,0	0,145 ± 0,036	0,681 ± 0,005
2022 деректерінде оқыту	One-Class SVM	2,0 / 2,0	0,374	0,739
2022 деректерінде оқыту	Автокодтаушы	1,2 ± 1,0 / 1,8 ± 0,4	0,361 ± 0,074	0,639 ± 0,053

Дереккөз: автордың ашық деректерді талдау нәтижелері (MetroPT2; тест 01.06–28.07.2022, 51,2 тәулік). Мәндер – 5 іске қосу бойынша орташа ± стандартты ауытқу.

Талқылау.

Нәтижелер ерте ескерту туралы күтуді бұл деректерде растамайды. MetroPT-3 деректерінде ауаның ағуы агрегат жұмысында есепте белгіленген уақытқа жақын байқалады, ал жалған дабылдар аз болуы талап етілгенде модельдер бұл ауытқуды кешігіп анықтайды немесе мүлде анықтамайды. Біздің хаттама нүктелік бағалау нәтижелерді асыра

көрсетуі мүмкін деген ескертулерді ескеріп құрылды (Kim және т.б., 2022: 7194–7201; Wu, Keogh, 2021): бүкіл тест кезеңі бағаланады, ал шекті мән алдын ала белгіленеді. Сонымен қатар есептегі басталу уақыты шартты екенін ескеру керек (мысалы, 1-ақау 00:00-ден басталады), сондықтан ондаған минуттық айырмалар нақты физикалық кешігуді білдірмеуі мүмкін.

Практикалық тұрғыдан маңызды нәтиже – жалған дабылдар құнының төмендеуі. Штаттық LPS ережесімен салыстырғанда модельдер мен EWMA картасы тәулігіне 0,5 деңгейде сол немесе одан да көп ақауды шамамен 3,4–4,4 есе аз жалған дабылмен анықтайды. Бұл ретте қарапайым EWMA картасы оқиғалар деңгейінде күрделі модельдерден кем түспейді, ал қатаң деңгейлерде олардан жақсы, демек кәсіпорында күйі бойынша қызмет көрсетуді енгізуді түсіндірілетін статистикалық бақылаудан бастауға болады, ал модельдерді терезелерді ажырату маңызды болған жерде қосуға болады.

Басқа кезеңге ауыстыру нәтижесі модель мен шекті мәнді бөлек қарастыру қажеттігін көрсетеді. Аномалиялардың реті басқа пайдалану кезеңінде сақталады, бірақ бағалардың абсолюттік деңгейі өзгереді, сондықтан шекті мәнді нақты нысанның деректерінде қайта калибрлеу керек. Бұл, әсіресе, нәтижелерді басқа өндірушінің немесе басқа пайдалану жағдайындағы өнеркәсіптік локомотивтер мен вагондардың пневматикалық жүйелеріне қолдану кезінде маңызды.

Зерттеудің шектеулері. Біріншіден, ақаулар саны аз (2020 жылы 4, 2022 жылы 2), сондықтан әдістер арасындағы оқиғалар деңгейіндегі айырмашылықтарды статистикалық тұрғыдан тексеру мүмкін емес; ең тұрақты нәтиже – жалған дабылдар санындағы айырмашылық (B0 үшін 75 эпизод). Екіншіден, ақау таңбалары (labels) – кәсіпорын есептеріндегі уақыт аралықтары, олар өлшенген эталон (ground truth) емес. Үшіншіден, нысан – метро пойызының агрегаты; нәтижелердің қазақстандық өнеркәсіптік көлік жабдығына жарамдылығы тексерілмеді. Төртіншіден, жарияланған деректер 0,1 Гц жиілікпен берілген, бастапқы 1 Гц жиілікте нәтиже өзгеше болуы мүмкін. Бесіншіден, LSTM сияқты рекуррентті модельдер қарастырылмады. Апаттылықтың азаюы туралы тұжырым жасалмайды: нақты пайдалану нәтижелері жоқ.

Қорытынды.

Зерттеу мақсатына қол жеткізілді: метро пойызының ауа өндіру агрегатының ашық IoT деректерінде аномалияларды анықтау модельдері (Isolation Forest, One-Class SVM, автокодтаушы) агрегаттың штаттық шекті ережесімен және EWMA бақылау картасымен оқиғалар деңгейінде салыстырылды. Салыстыру хронологиялық хаттама бойынша жүргізілді: модельдер ақпан айының деректерінде оқытылды, шекті мән наурыз айында берілген тәуліктік жалған дабыл деңгейіне калибрленді, ал сапа сәуір–тамыз аралығында ақау есептерімен салыстырылып бағаланды. Нәтиженің тұрақтылығы басқа пайдалану кезеңінің (2022 ж.) деректерінде тексерілді.

Алынған нәтижелер (ашық деректерді талдау). Біріншіден, алдын ала белгіленген тәулігіне 0,1 және 0,2 жалған дабыл деңгейлерінде ешбір әдіс ақауды есептегі басталу уақытынан бұрын анықтаған жоқ, ескерту уақытының медианасы барлық деңгейлерде теріс болды; сондықтан H1 болжамы расталмады. Екіншіден, тәулігіне 0,5 деңгейге калибрленгенде модельдер мен EWMA орта есеппен 4 ақаудың 3,8–4-ін тестте тәулігіне 0,15–0,19 жалған дабылмен анықтады, ал штаттық ереже 3 ақауды тәулігіне 0,66 жалған дабылмен анықтады; яғни жалған дабылдар 3,4–4,4 есе аз. Үшіншіден, модельдер ақаулар маңындағы терезелерді EWMA картасына қарағанда жақсырақ ажыратады (PR-AUC 0,40–0,45 және 0,24, кездейсоқ деңгей 0,056), бірақ оқиғалар деңгейінде одан озбайды, ал қатаң шекті мәндерде EWMA модельдерден де жақсы болды. Төртіншіден, 2020 жылғы модельдердің ажырату қабілеті 2022 жылғы деректерде сақталды (PR-AUC 0,69–0,73), бірақ 2020 жылғы шекті мән жарамсыз болды: тәулігіне 0,1 деңгейде бірде-бір ақау анықталмады, ал 0,5 деңгейде ақаулар басталғаннан 15–21 сағ кейін ғана анықталды; сондықтан H2

ішінара расталды. 2022 жылғы деректерде қайта оқыту екі ақауды да анықтауға мүмкіндік берді, бірақ бұл нәтиже тек екі оқиғаға негізделген және дәлел ретінде қарастырылмайды.

Ғылыми жаңалығы – метро компрессорының нақты деректерінде аномалияларды анықтау модельдерін агрегаттың өзінің штаттық ережесімен және қарапайым статистикалық картамен бірдей хаттама бойынша, оқиғалар деңгейінде (ақауды анықтау, ескерту уақыты, тәулігіне жалған дабылдар) салыстыру және модель мен шекті мәннің басқа кезеңге ауысуын бөлек бағалау. Бұл нәтижелер терезелер деңгейіндегі жоғары метрикалар оқиғалар деңгейіндегі ерте ескертуге автоматты түрде айналмайтынын көрсетеді.

Әдістемелік қорытынды. Мұндай жүйелерді бағалау кезінде (1) деректерді уақыт бойынша бөлу және шекті мәнді тест кезеңіне дейін бекіту, (2) нәтижелерді терезелер деңгейінде ғана емес, оқиғалар деңгейінде – анықталған ақаулар саны, ескерту уақыты және тәулігіне жалған дабылдар саны бойынша – көрсету, (3) модельдерді жабдықтың штаттық ережесімен және қарапайым статистикалық бақылау картасымен міндетті түрде салыстыру қажет. Сонымен қатар ақау таңбаларының сапасын тексеру керек: тест кезеңінде есептерде жоқ, шамамен 62 сағатқа созылған үздіксіз жүктеме эпизоды анықталды, оны модельдер аномалия ретінде белгіледі, ал бағалауда ол жалған дабыл болып есептелді.

Практикалық маңызы. Өнеркәсіптік көлік кәсіпорнында күйі бойынша техникалық қызмет көрсетуді енгізу кезінде дабыл шегін техникалық қызмет көрсету қызметі қабылдай алатын тәуліктік жалған дабылдар санына сүйеніп таңдау және оны әр нысан мен әр пайдалану кезеңі үшін қайта калибрлеу ұсынылады. Енгізуді түсіндірілетін EWMA картасынан бастауға, ал модельдерді ақаулар маңындағы аралықтарды ажырату маңызды болған жерде қосуға болады. Нәтижелерді өнеркәсіптік локомотивтер мен вагондардың пневматикалық жүйелеріне қолдану үшін кәсіпорында қысым, температура, қозғалтқыш тогы және компрессордың жұмыс күйі сигналдарын тұрақты жиілікпен жинайтын телеметрия, сондай-ақ ақаулар мен жөндеулердің уақыт белгісі бар журналы болуы қажет; онсыз шекті мәнді калибрлеу мен тексеру мүмкін емес. Дабылдар CMMS жүйесіне тексеру тапсырмасы ретінде берілуі мүмкін; бұл ретте нәтижелер апаттылықтың азаюы туралы тұжырым жасауға негіз бермейді, себебі нақты пайдалану нәтижелері жоқ.

Шектеулер мен болашақ жұмыс. Ақаулар саны аз (2020 жылы 4, 2022 жылы 2), ақау таңбалары кәсіпорын есептеріндегі уақыт аралықтары ғана, ал нысан – метро пойызының агрегаты. Сондықтан келесі кезеңде қазақстандық өнеркәсіптік көлік кәсіпорнының жабдығы бойынша деректер мен ақаулар журналын алу, бастапқы 1 Гц жиіліктегі деректермен тексеру, рекуррентті модельдерді қосу және дабылдардың техникалық қызмет көрсету шешімдеріне әсерін нақты пайдалануда бағалау жоспарланады.

REFERENCES

- Barros M., Veloso B., Pereira P.M., Ribeiro R.P., Gama J. (2020). Failure Detection of an Air Production Unit in Operational Context // *Communications in Computer and Information Science*. Pp. 61–74. https://doi.org/10.1007/978-3-030-66770-2_5
- Carvalho T.P., Soares F.A.A.M.N., Vita R., Francisco R. da P., Basto J.P., Alcalá S.G.S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance // *Computers & Industrial Engineering*. Vol. 137. — Art. 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- Davari N., Veloso B., Ribeiro R.P., Pereira P.M., Gama J. (2021). Predictive maintenance based on anomaly detection using deep learning for air production unit in the railway industry // *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*. Pp. 1–10. <https://doi.org/10.1109/DSAA53316.2021.9564181>
- Davari N., Veloso B., Ribeiro R., Gama J. (2023). MetroPT-3 Dataset [Dataset] // *UCI Machine Learning Repository*. <https://doi.org/10.24432/C5VW3R>
- Davari N., Veloso B., Ribeiro R.P., Gama J. (2024). Detecting and Explaining Anomalies in the Air Production Unit of a Train // *Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing*. — Pp. 358–364. <https://doi.org/10.1145/3605098.3635906>
- Davis J., Goadrich M. (2006). The relationship between Precision-Recall and ROC curves // *Proceedings of the 23rd International Conference on Machine Learning*. Pp. 233–240. <https://doi.org/10.1145/1143844.1143874>

- García-Méndez S., de Arriba-Pérez F., Leal F., Veloso B., Malheiro B., Burguillo-Rial J.C. (2025). An explainable machine learning framework for railway predictive maintenance using data streams from the metro operator of Portugal // *Scientific Reports*. Vol. 15. — Art. 27495. <https://doi.org/10.1038/s41598-025-08084-1>
- Hundman K., Constantinou V., Laporte C., Colwell I., Soderstrom T. (2018). Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding // *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Pp. 387–395. <https://doi.org/10.1145/3219819.3219845>
- Kim S., Choi K., Choi H.-S., Lee B., Yoon S. (2022). Towards a Rigorous Evaluation of Time-Series Anomaly Detection // *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36, No. 7. — Pp. 7194–7201. <https://doi.org/10.1609/aaai.v36i7.20680>
- Liu F.T., Ting K.M., Zhou Z.-H. (2008). Isolation Forest // *2008 Eighth IEEE International Conference on Data Mining*. Pp. 413–422. <https://doi.org/10.1109/ICDM.2008.17>
- Meira J., Veloso B., Bolón-Canedo V., Marreiros G., Alonso-Betanzos A., Gama J. (2023). Data-driven predictive maintenance framework for railway systems // *Intelligent Data Analysis*. Vol. 27. No. 4. — Pp. 1087–1102. <https://doi.org/10.3233/IDA-226811>
- Roberts S.W. (1959). Control Chart Tests Based on Geometric Moving Averages // *Technometrics*. Vol. 1, No. 3. — Pp. 239–250. <https://doi.org/10.1080/00401706.1959.10489860>
- Saito T., Rehmsmeier M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets // *PLOS ONE*. Vol. 10. No. 3. — Art. e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Schölkopf B., Platt J.C., Shawe-Taylor J., Smola A.J., Williamson R.C. (2001). Estimating the Support of a High-Dimensional Distribution // *Neural Computation*. Vol. 13. No. 7. — Pp. 1443–1471. <https://doi.org/10.1162/089976601750264965>
- Veloso B., Ribeiro R.P., Gama J., Pereira P.M. (2022). The MetroPT dataset for predictive maintenance // *Scientific Data*. Vol. 9. — Art. 764. <https://doi.org/10.1038/s41597-022-01877-3>
- Veloso B., Gama J., Ribeiro R., Pereira P. (2022). MetroPT2: A Benchmark dataset for predictive maintenance [Dataset], Version V2 // *Zenodo*. <https://doi.org/10.5281/zenodo.7766691>
- Wu R., Keogh E. (2021). Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress // *IEEE Transactions on Knowledge and Data Engineering*. (early access). <https://doi.org/10.1109/TKDE.2021.3112126>
- Zafra A., Veloso B., Gama J. (2024). Early Failure Detection for Air Production Unit in Metro Trains // *Lecture Notes in Computer Science*. Pp. 339–351. https://doi.org/10.1007/978-3-031-74183-8_28
- Zio E. (2022). Prognostics and Health Management (PHM): Where are we and where do we (need to) go in theory and practice // *Reliability Engineering & System Safety*. Vol. 218. — Art. 108119. <https://doi.org/10.1016/j.ress.2021.108119>