|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Государственная корпорация по атомной энергии «Росатом» | | | | | | |
| Организация  УДК:  Инв. №: | | | | | | |
|  | | | | УТВЕРЖДАЮ | | |
|  | | | | |  | | --- | | Должность | | | |
|  | |  | |  | | Ф.И.О. |
|  | | | | «\_\_\_» \_\_\_\_\_\_\_\_\_ 2024 г. | | |
| Номер документа | | | | | | |
| ОТЧЕТ О проверке гипотезы: | | | | | | |
| Бинарная классификация аномального режима работы насосного оборудования на данных о параметрах его работы | | | | | | |
|  | | | | | | |
| Договор № Курс ИИ | | | | | | |
|  | | | | | | |
|  | Руководитель проекта: | | Шадрина О.Е. | | Ф.И.О.  2024 г. | | |
| « 01 » 09 | |
|  | |  | | |
|  |  | |  | | | | |
|  | Москва 2024 | | | | | | |

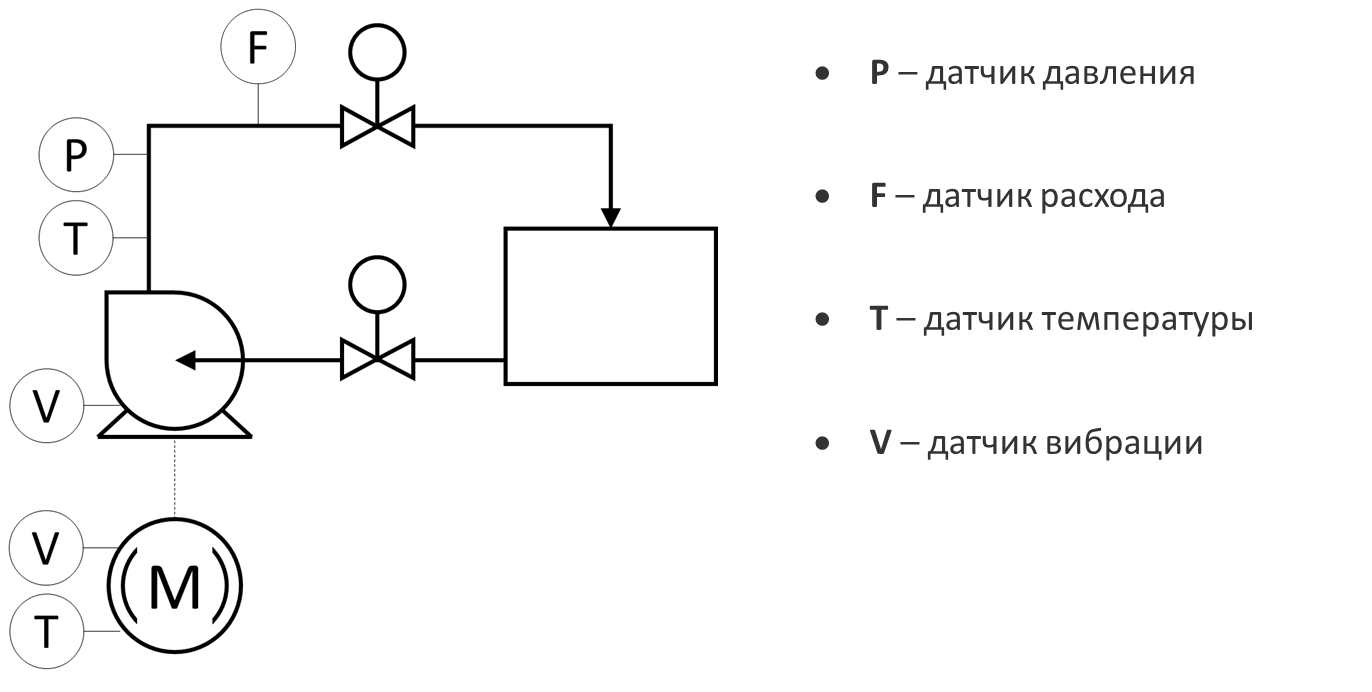
1. **Наименование работ**

Работы по проверке гипотезы «Бинарная классификация аномального режима работы оборудования на данных о параметрах его работы».

1. **Описание процесса**

Технологический процесс остужения теплоносителя в конденсаторе заключается в прокачке хладагента по внутритрубному пространству конденсатора. Хладагент забирается из резервуара, прокачивается горизонтальным центробежным насосом через конденсатор и возвращается в резервуар.

Упрощенно схема процесса представлена на рисунке ниже:



1. **Гипотеза**

Бинарная классификация аномального режима работы оборудования на данных о параметрах его работы.

1. **Критерий успешности**

Baseline побит (F1> 0.84464)

1. **Метрики**

F1

1. **Исходные данные**

Исходные данные представляют из себя многомерный временной ряд, собранный с датчиков технологической линии.

Перечень зарегистрированных параметров представлен в таблице ниже:

|  |  |  |
| --- | --- | --- |
| **Параметр** | **Описание** | **Единицы измерения** |
| Accelerometer1RMS | Виброускорение | - |
| Accelerometer2RMS | Виброускорение | - |
| Current | Сила тока питания электродвигателя | А |
| Pressure | Давление на выкиде | Bar |
| Temperature | Температура корпуса электродвигателя | oC |
| Thermocouple | Температура перекачиваемой среды (воды) | oC |
| Voltage | Напряжение питания электродвигателя | В |
| RateRMS | Расход перекачиваемой среды (воды) | л/мин |
| anomaly | Флаг аномалии (дискретный параметр) | - |
| changepoint | Флаг изменения состояния (дискретный параметр) | - |

Имеются значения указанных параметров за отдельные периоды измерении в 2023 году.

Дискретизация параметров составляет для большинства данных – 1 секунда с периодическими пропусками.

1. **Модели**

Подготовка данных:

* Приведены к 0 все отрицательные параметры измерений
* Сгенерированы признаки среднего и разности виброускорений, температур
* Сгенерированы признаки «сглаживания» измерений
* Проведено удаление выбросов
* Объединены все временные ряды

Для проверки гипотез был использован преподавательский подход с автоматическим подбором параметров.

Далее были использованы наиболее популярные нейронные сети: «Случайный лес», Сверточный автокодер TensorFlow для одномерных временных рядов

При выборе оптимальных параметров для нейросети «Случайный лес» отдельно для каждого временного ряда была использована библиотека Optuna, применение которой показало неоднородность временных рядов и, как результат, возможность достижения наименьшей средней квадратической ошибки (MSE) на каждом временном ряду только при индивидуальных параметрах сети.

Так здесь наихудшие результаты (MSE=0,12) даёт временной ряд /data/other/13.csv. Лучшие (MSE=0,02) даёт временной ряд / data/valve1/12.csv. Остальные временные ряды показывают результат в MSE<=0.08.

Предположительно временной ряд /data/other/13.csv, а также разница в уровнях измерений по рядам не дают возможности привести объединённый временной ряд к оптимальному виду, что не даёт существенного увеличения точности.

**Результаты**

Первый вариант с преподавательским подходом дал результат F1= 0.84464, далее был получен результат F1= 0.84355.

Данный результат свидетельствует о необходимости изменения в методологии подготовки данных (выравнивание рядов, деление аномалий по типам аномалий для разных рядов), что не удалось сделать в условиях недостатка времени.

1. **Выводы**

Поскольку не удалось улучшить результат baseline, гипотеза о бинарной классификации аномального режима работы оборудования на данных о параметрах его работы считается подтверждённой на уровне F1= 0.84464.