

Постановка задачи

В процессе эксплуатации промышленных электродвигателей возникают различные дефекты, вызванные износом компонентов, усталостью материалов или неправильной эксплуатацией. Целью данной работы является классификация таких дефектов на основе анализа временных рядов электрических сигналов, что позволит обнаружить и предотвратить отказ на ранних стадиях. Датасет включает в себя временные ряды электрических сигналов тока и напряжения для трех фаз каждого двигателя, как в здоровом состоянии, так и при наличии дефектов, таких как замыкание витков статора, обрыв стержня ротора, износ подшипников и другие. Замеры по 12 с и делились на сэмплы по 0,5 с. Далее из каждого сэмпла извлекали признаки такие как: минимальное, максимальное, среднее, среднеквадратичное значения, а также использовались Фурье, вейвлет анализ и некоторые другие методы.

Исследуемые модели машинного обучения

В рамках исследования было разработано и протестировано три модели для решения задачи классификации дефектов промышленных электродвигателей.

1. Autoencoder CNN

Задача: бинарная классификация (здоровый двигатель и дефект подшипника).

Архитектура: автоэнкодер, использующий сверточные слои (Conv1D) для извлечения признаков из временных рядов.

Особенности: после обработки данных модель обучалась на сэмплах по 0,5 секунд, с 174 признаками до преобразования и 407 признаками после применения преобразования Парка. В качестве функции активации использовалась ReLU, оптимизатор — Adam, функция потерь — среднеквадратичная ошибка (MSE).

Результаты: Модель показала точность 0.44 на тестовых данных, с F1-мерой для класса 1 (дефект подшипника) на уровне 0.59, что указывает на потребность в дальнейшей доработке модели.

True	0	19	251
	1	154	446
		0	1
	Predicted		

2. CatBoost (Модель 2)

Задача: Бинарная классификация (здоровый двигатель и обрыв стержня ротора).

Архитектура: Классическая модель градиентного бустинга на деревьях решений (CatBoost).

Особенности: Датасет был предварительно обработан, данные были преобразованы в сэмплы с 407 признаками после применения преобразования Парка.

Результаты: Точность модели составила 0.81 на тестовых данных, F1-мера для класса 1 составила 0.43.

True	0	302164	45836
	1	30693	29307
		0	1
	Predicted		

	precision	recall	F1-score	support
0	0.91	0.87	0.89	348 000
1	0.39	0.49	0.43	60 000
accuracy			0.81	408 000
macro avg	0.65	0.68	0.66	408 000
weighted avg	0.83	0.81	0.82	408 000

3. FEDOT AutoML (Модель 3)

Задача: Множественная классификация (здоровый двигатель и различные дефекты подшипников).

Архитектура: Использование AutoML платформы FEDOT с настройкой автоматического подбора лучших моделей и параметров.

Особенности: Датасет содержал сэмплы с 407 признаками после преобразования Парка. Модель использовала комбинацию различных методов, включая CatBoost и Random Forest.

Результаты: Модель достигла точности 0.84 на тестовых данных, с F1-мерой для класса 1 (дефект подшипника) на уровне 0.19 и более высокими показателями для других классов.

	precision	recall	F1-score	support
accuracy			0.59	2670
macro avg	0.62	0.59	0.58	2670
weighted avg	0.60	0.59	0.57	2670

	F1-score
> 0.8	здоровый, обрыв стержня ротора, дефект рабочего колеса насоса, несоосность насоса и двигателя, дефекты подшипников
0.5 – 0.8	загрязнение подшипника, дефект крепления двигателя, деформация опоры, межвитковые замыкания статора
< 0.5	кавитация, сцепление станины с опорой, деформация муфты, разбалансировка двигателя и насоса

Основные выводы

1. Решение задачи классификации дефектов на основе временных рядов электрических сигналов позволяет заранее обнаружить неисправности и предотвратить внезапные отказы оборудования, что критично для промышленных предприятий.

2. Модель FEDOT AutoML показала наилучшие результаты в задаче множественной классификации, однако модель демонстрирует недостаточную точность распознавания дефектов на начальных стадиях их развития.

3. В дальнейшем планируется сформировать комплексную модель, которая бы на первом этапе решала задачу бинарной классификации наличия дефекта, а на втором – задачу множественной классификации тип дефекта. Кроме того, требует решения проблема дисбалансов классов и правильной обработки данных перед обучением модели. Отдельного рассмотрения заслуживает исследование, какие признаки более важны при обучении.

4. В ходе работы возникла гипотеза о том, что дополнение датасета данными с датчиков вибрации позволит улучшить результаты работы моделей.