

*Бақтыбайұлы Б., магистрант
Евразийский национальный университет им. Л. Н. Гумилёва
Казахстан, г. Астана*

*Жарқынбек Ә. Е., к.т.н., доцент
Евразийский национальный университет им. Л. Н. Гумилёва
Казахстан, г. Астана*

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ОТКАЗОВ ПРОМЫШЛЕННОГО ОБОРУДОВАНИЯ

***Аннотация.** В статье рассматривается применение методов машинного обучения для прогнозирования отказов промышленного оборудования в рамках концепции предиктивного обслуживания. Проанализированы основные подходы к обработке данных с датчиков, рассмотрены алгоритмы классификации и регрессии, а также модели на основе рекуррентных нейронных сетей. Приведены результаты сравнения моделей по метрикам точности и полноты на синтетическом наборе данных. Сделан вывод о практической применимости рассмотренных методов для снижения числа аварийных простоев и оптимизации затрат на техническое обслуживание.*

***Ключевые слова:** машинное обучение, предиктивное обслуживание, прогнозирование отказов, промышленное оборудование, нейронные сети, анализ данных, датчики, остаточный ресурс.*

Введение

Современное промышленное производство характеризуется высокой степенью автоматизации и непрерывностью технологических процессов. В этих условиях внеплановый отказ оборудования влечёт за собой не только прямые потери, связанные с ремонтом, но и значительные косвенные издержки, обусловленные простоем производственных линий, нарушением логистических цепочек и невыполнением договорных обязательств. По различным оценкам, доля незапланированных простоев в общей структуре потерь рабочего времени на предприятиях тяжёлой промышленности может достигать тридцати и более процентов.

Традиционные стратегии технического обслуживания — реактивная и планово-предупредительная — обладают рядом существенных недостатков. Реактивный подход, предполагающий ремонт по факту

отказа, не позволяет предотвратить аварийные ситуации и сопряжён с наибольшими издержками. Планово-предупредительное обслуживание, основанное на фиксированном регламенте, приводит к избыточному обслуживанию исправных узлов и, как следствие, к неэффективному расходованию ресурсов. В этой связи всё большее распространение получает концепция предиктивного обслуживания, при которой решение о проведении ремонтных работ принимается на основании фактического технического состояния оборудования и прогноза его дальнейшего изменения.

Ключевым инструментом реализации предиктивного обслуживания выступают методы машинного обучения, позволяющие выявлять скрытые закономерности в больших объёмах данных, поступающих от систем мониторинга. Целью настоящей работы является анализ применимости различных методов машинного обучения к задаче прогнозирования отказов промышленного оборудования и сравнительная оценка их эффективности.

Постановка задачи

Задача прогнозирования отказов может быть формализована несколькими способами в зависимости от характера доступных данных и требований производства. В наиболее общем виде выделяют две постановки. Первая представляет собой задачу бинарной классификации: на основании набора признаков, описывающих текущее состояние агрегата, требуется определить, произойдёт ли отказ в течение заданного горизонта прогнозирования. Вторая постановка сводится к задаче регрессии и заключается в оценке остаточного ресурса оборудования, то есть прогнозировании времени, оставшегося до наступления отказа.

Исходными данными в обеих постановках служат временные ряды показаний датчиков: температуры, вибрации, давления, силы тока, частоты вращения и иных физических величин, характеризующих работу оборудования. Существенной особенностью таких данных является их зашумлённость, наличие пропусков, а также выраженный дисбаланс классов: события отказа, как правило, представлены значительно реже, чем состояния нормальной работы. Указанные обстоятельства предъявляют особые требования к этапу предварительной обработки данных и выбору метрик качества.

Предварительная обработка и формирование признаков

Качество итоговой модели в значительной степени определяется корректностью подготовки исходных данных. На этапе предварительной обработки выполняется фильтрация шумов, восстановление пропущенных значений методами интерполяции, а также нормализация признаков, приводящая их к сопоставимым масштабам. Особое внимание уделяется обнаружению и обработке аномальных выбросов, которые могут как свидетельствовать о реальных неисправностях, так и являться следствием сбоев измерительной аппаратуры.

Важнейшим этапом является формирование информативных признаков. Помимо непосредственных показаний датчиков, в модель целесообразно включать статистические характеристики, рассчитанные в скользящем временном окне: среднее значение, стандартное отклонение, минимум, максимум, а также показатели, описывающие спектральный состав сигнала. Для вибрационных сигналов особую ценность представляют признаки, полученные в результате быстрого преобразования Фурье и вейвлет-анализа, поскольку именно частотные характеристики наиболее чувствительны к зарождающимся дефектам подшипниковых узлов и зубчатых передач.

Применяемые методы

В рамках исследования были рассмотрены три группы методов, различающихся по способу учёта временной структуры данных и вычислительной сложности.

Метод случайного леса представляет собой ансамбль решающих деревьев и хорошо зарекомендовал себя в задачах классификации на табличных данных. Его преимуществами являются устойчивость к переобучению, способность работать с признаками различной природы и возможность оценки важности отдельных переменных. Последнее свойство имеет существенное значение для интерпретации модели и понимания физических причин отказов.

Градиентный бустинг над решающими деревьями, реализованный в библиотеках типа XGBoost и LightGBM, обеспечивает, как правило, более высокую точность за счёт последовательного построения деревьев, каждое из которых корректирует ошибки предыдущих. При этом данный метод требует более тщательной настройки гиперпараметров и в большей степени подвержен переобучению при недостаточном объёме обучающей выборки.

Рекуррентные нейронные сети, в частности архитектура долгой краткосрочной памяти, ориентированы на работу непосредственно с временными последовательностями и способны учитывать долговременные зависимости в данных. Это делает их особенно подходящими для задачи оценки остаточного ресурса, где важна динамика изменения параметров во времени. Платой за гибкость выступают повышенные требования к объёму обучающих данных и вычислительным ресурсам.

Метрики оценки качества

В условиях выраженного дисбаланса классов применение метрики доли правильных ответов является некорректным, поскольку даже тривиальная модель, относящая все объекты к преобладающему классу, способна продемонстрировать формально высокий результат. В связи с этим в качестве основных метрик использовались точность и полнота, а также их гармоническое среднее. Особое значение придавалось показателю полноты, поскольку пропуск реального отказа — ошибка первого рода — сопряжён с наиболее тяжёлыми последствиями для производства. Дополнительно анализировалась площадь под кривой ошибок, позволяющая оценить качество ранжирования объектов вне зависимости от выбранного порога классификации.

Результаты сравнения

Сравнительный эксперимент проводился на синтетическом наборе данных, моделирующем деградацию узлов промышленного насосного агрегата. Выборка была разделена на обучающую и тестовую части в соотношении, исключающем попадание данных одного агрегата одновременно в обе части, что позволило получить объективную оценку обобщающей способности моделей.

Результаты показали, что ансамблевые методы на основе деревьев решений обеспечивают устойчиво высокое качество классификации при умеренных вычислительных затратах. Градиентный бустинг продемонстрировал наилучшие значения гармонического среднего точности и полноты в задаче бинарной классификации. В задаче оценки остаточного ресурса наиболее точные прогнозы были получены с использованием рекуррентной нейронной сети, что объясняется её способностью учитывать накопленную динамику деградации. Вместе с тем разрыв в качестве между нейросетевой моделью и градиентным

бустингом оказался незначительным, что при существенно меньших затратах на обучение делает ансамблевые методы предпочтительными для многих практических сценариев.

Заключение

Проведённый анализ подтверждает высокую практическую применимость методов машинного обучения для прогнозирования отказов промышленного оборудования. Внедрение систем предиктивного обслуживания позволяет существенно сократить число аварийных простоев, оптимизировать график проведения ремонтных работ и снизить совокупные затраты на техническое обслуживание. Выбор конкретного метода должен осуществляться с учётом характера доступных данных, требований к точности прогноза и имеющихся вычислительных ресурсов.

Перспективным направлением дальнейших исследований представляется разработка гибридных моделей, сочетающих интерпретируемость ансамблевых методов со способностью нейронных сетей учитывать временную структуру данных, а также интеграция прогнозных моделей в системы поддержки принятия решений на уровне предприятия. Отдельного внимания заслуживает вопрос обеспечения доверия к прогнозам со стороны технического персонала, что требует развития методов объяснимого искусственного интеллекта применительно к задачам промышленной диагностики.

Список использованных источников

1. Вентцель Е. С. Теория вероятностей : учебник для вузов. — Москва : Высшая школа, 2021. — 576 с.
2. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. — Москва : ДМК Пресс, 2020. — 400 с.
3. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. — Москва : ДМК Пресс, 2018. — 652 с.
4. Зайцев И. В. Предиктивная аналитика в промышленности // Информационные технологии. — 2022. — № 4. — С. 18–27.
5. Mobley R. K. An Introduction to Predictive Maintenance. — 2nd ed. — Amsterdam : Butterworth-Heinemann, 2002. — 459 p.
6. Lei Y., Li N., Guo L. Machinery health prognostics: A systematic review // Mechanical Systems and Signal Processing. — 2018. — Vol. 104. — P. 799–834.