

Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет»

На правах рукописи

Гончаров Аркадий Сергеевич

**АВТОМАТИЗАЦИЯ УПРАВЛЕНИЯ ПРЕДИКТИВНЫМ ТЕХНИЧЕСКИМ
ОБСЛУЖИВАНИЕМ И РЕМОНТОМ ПРОМЫШЛЕННЫХ РОБОТОВ**

Специальность 2.3.3 — Автоматизация и управление технологическими
процессами и производствами

Диссертация на соискание ученой степени кандидата технических наук

Научный руководитель:
канд. техн. наук, доцент
Савельев Алексей Олегович

Томск 2024

Оглавление

ВВЕДЕНИЕ.....	5
1. Актуальные подходы и методы решения задач оценки работоспособности и состояния промышленных роботов	13
1.1.Цифровизация производства как элемент автоматизированного управления технологическими процессами.....	13
1.2.Подходы и методы интеллектуализации процессов управления технологическими процессами и производствами	18
1.3.Подходы и ограничения процессов интеллектуальной оценки состояния промышленных роботов.....	31
1.4.Выводы по ГЛАВЕ 1.....	44
2. Проблемно-ориентированный конвейер обработки данных для прогнозирования работоспособности промышленных роботов с использованием методов интеллектуального анализа.....	46
2.1.Формальное представление процесса принятия решений ТОиР в рамках предиктивного подхода	46
2.2.Методика проектирования проблемно-ориентированного конвейера обработки данных, обеспечивающего реализацию процесса принятия решений ТОиР на основе данных.....	48
2.3.Подходы и методы управления моделями анализа данных	55
2.4.Методы оценки сходства данных.....	66
2.5.Методы оценки экономической эффективности и принятия решений.....	68
2.6.Проектирование комплекса алгоритмов для проблемно-ориентированного конвейера обработки данных.....	72
2.7.Выводы по ГЛАВЕ 2.....	78
3. Программное обеспечение автоматизации процессов ТОиР, реализующего разработанный комплекс алгоритмов	79
3.1.Проектирование архитектуры программной системы	79
3.1.1. Описание архитектуры на основе выявленных ограничений	79

3.1.2. Диаграмма компонентов программного обеспечения	81
3.2. Модуль управления данными	85
3.2.1. Диаграмма вариантов использования модуля управления данными	85
3.2.2. База данных хранения исторических данных о работе оборудования (на примере промышленных роботов)	86
3.2.3. Модуль выявления аномальных значений	88
3.2.4. Алгоритм оценки сходства данных	91
3.3. Модуль управления моделями анализа данных	93
3.3.1. Диаграмма вариантов использования модуля анализа данных	93
3.3.2. Алгоритм обучения и тестирования модели анализа данных	94
3.3.3. Алгоритм подбора модели машинного обучения на основе оценки сходства данных	95
3.3.4. Диаграммы последовательности модуля анализа данных	100
3.3.5. База данных для хранения моделей интеллектуального анализа	105
3.4. Модуль поддержки принятия решений на основе анализа данных	106
3.5. Выводы по ГЛАВЕ 3	109
4. Апробация программного обеспечения автоматизации процессов ТОиР с применением методов интеллектуального анализа данных	110
4.1. Характеристики наборов данных, использованных для апробации	110
4.2. Подготовка данных для апробации	112
4.3. Результаты применения алгоритма оценки сходства сформированных выборок	115
4.4. Обучение и тестирование базовых моделей	116
4.5. Тестирование моделей на выборках, сформированных с парка роботов ..	117
4.6. Оценки эффективности применения разработанного программного обеспечения	123
4.7. Разработка имитационных моделей для расчета экономической эффективности	127
4.8. Внедрение результатов диссертационного исследования	133
4.9. Выводы по ГЛАВЕ 4	134

Заключение	135
Список сокращений и условных обозначений.....	137
Список литературы	138
Приложение 1 – Публикации по теме исследования	156
Приложение 2 – Результаты интеллектуальной деятельности	158
Приложение 3 – Диаграммы процесса принятия решений по ТОиР на основе анализа данных (в нотации IDEF0)	160
Приложение 4 – Логическая схема базы данных для хранения исторических данных о техническом обслуживании промышленного оборудования	164
Приложение 5 – Алгоритм обучения и тестирования модели на наборах данных разнородного оборудования.....	165
Приложение 6 – Алгоритмы обучения и тестирования модели анализа данных	170
Приложение 7 – Структура файла для апробации алгоритмической модели ...	172
Приложение 8 – Комбинации признаков для отбора значимых параметров	173
Приложение 9 – Результаты нормализации датасетов.....	175
Приложение 10 – Графики зависимости точности моделей машинного обучения от сходства датасетов на примере задачи прогнозирования температуры привода	178
Приложение 11 – Алгоритм детектирования аномальных значений (в нотации EPC)	182
Приложение 12 – Акты внедрения результатов диссертационной работы	183

ВВЕДЕНИЕ

Актуальность темы исследования. Оценка текущего состояния оборудования в составе сложных технологических комплексов и прогнозирования его работоспособности в настоящее время является одной из приоритетных задач построения современного «интеллектуального» производства. Проблема управления промышленным оборудованием является комплексной, включающей технический, технологический, управленческий, экономический и организационный аспекты. Решение такого рода проблем требует разработки методов и алгоритмов автоматизации процессов мониторинга и прогнозирования технического состояния оборудования в реальном времени. Обслуживание современных сложных технических систем требует значительных материальных затрат. Актуальной остается задача по обеспечению безотказного функционирования промышленного оборудования при минимизации затрат на его содержание и техническое обслуживание. Принятие решений при техническом обслуживании основывается на извлечении и получении новых знаний, добываемых из низкоуровневых данных. Современные исследования по применению интеллектуальных методов анализа данных для решения задач прогнозирования работоспособности и отказов промышленного оборудования лежат в рамках подходов по проактивному и предиктивному техническому обслуживанию. Данные подходы позволяют осуществить переход от реактивного и планового обслуживания к обслуживанию на основе данных о прогнозном состоянии или полезном остаточном рабочем ресурсе оборудования.

Разработка алгоритмов автоматизации управления техническим состоянием оборудования, прогнозирования работоспособности и поддержки принятия решений для технического обслуживания и ремонта на основе данных является актуальной задачей. Разработка технологий интеллектуализации производственных процессов и методов прогнозной аналитики является важной задачей для эффективного анализа данных и извлечения полезных знаний для решения соответствующих практических задач.

Степень изученности темы исследования. Научные направления по изучению систем и процессов технического обслуживания в различных отраслях промышленности рассматривались такими российскими и зарубежными авторами как: А.А. Мисловидова, И.О. Бельский, К.С. Ахмедов, А.А. Сычугов, И.В. Сорокина, А.Е. Савочкин и др. Вклад в исследования по разработке и реализации методов интеллектуального анализа данных для процесса технического обслуживания внесли: И.С. Макаш, С.В. Квонг, А.В. Суханов, М.Ш. Омар, И.П. Врублевский, О.Н. Долинина, Е.Е. Дышкант, А.В. Кизим, L. Zhang, C.L. Wang, E. Sezer, D. Romero, F. Guedea, R. Chianese, D.K. Saha, S. Ahmed и др. Исследованию и разработке методов прогнозирования работоспособности, ранних отказов и остаточного ресурса промышленного оборудования уделяли внимание следующие ученые: P. Singh, S. Agrawal, J. Dsouza, S. Velan, R. Ashmore, L. Zhang, R. Calinescu, L.P. Silvestrin, G. Ortiz и др. Вызовы и проблемы использования и управлением моделями машинного обучения и большими данными исследовались: S. Schelter, F. Biessmann, T. Januschowski, F. McSherry, A. Kumar, R. J Hyndman и др. Многие современные отечественные (ООО «Цифра», Ростех, Северсталь, МАК «Вымпел» и др.) и международные (SAP, AT&T, Siemens, General Electric, Microsoft, KUKA, ABB, Fanuc и др.) компании также развивают направления по предиктивному техническому обслуживанию и принятию решений на основе данных. Наличие теоретических исследований и практических решений, внедренных в производственные циклы, свидетельствует об актуальности данного направления и высокой степени его разработанности. Тем не менее прогресс в области разработки эффективных методов выбора оптимальных прогнозных моделей интеллектуального анализа данных предоставляет широкое поле для изучения, что определяет необходимость выполнения дополнительных исследований.

Объект исследования – процесс управления предиктивным техническим обслуживанием и ремонтом промышленных роботов.

Предмет исследования – алгоритмическое и программное обеспечение для автоматизации управления предиктивным техническим обслуживанием и ремонтом промышленных роботов.

Цель работы – повышение эффективности процесса автоматизированного управления техническим обслуживанием и ремонтом промышленных роботов с применением комплекса алгоритмов управления моделями анализа данных.

Задачи

- Анализ изученности темы проактивного и предиктивного обслуживания промышленного оборудования с использованием методов анализа данных.
- Разработка архитектуры программного обеспечения с использованием конвейерного подхода, допускающего возможность расширения, тиражирования и адаптивности под специфику конкретного производственного процесса для обеспечения стратегического управления и планирования технического обслуживания парка разнородного производственного оборудования (на примере промышленных роботов).
- Разработка комплекса алгоритмов управления и выбора эффективных прогнозных моделей для анализа данных разнородного промышленного оборудования с учетом настройки гиперпараметров моделей машинного обучения, а также минимизацией вычислительных и временных ресурсов на обучение моделей.
- Разработка и апробация программного обеспечения для исследования процесса анализа данных промышленного робота с использованием методов анализа данных и оценки сходства наборов данных.

Методы исследования. Для решения поставленных задач были использованы следующие методы: имитационное моделирование, алгоритмы машинного обучения и искусственных нейронных сетей, методы оценки сходства данных, системный анализ.

Научная новизна

- Предложена методика проектирования информационной системы для автоматизации предиктивного технического обслуживания и ремонта

промышленных роботов, отличающаяся обеспечением оркестрации, модульности и масштабируемости на всех этапах технического обслуживания.

- Разработан алгоритм автоматизированного подбора и оценки эффективности модели машинного обучения для прогнозирования параметров работы промышленного робота, отличающийся сокращением времени оценки его состояния до 18%.

- Разработан алгоритм вычисления меры сходства наборов исторических данных о работе промышленного робота, повышающий эффективность процесса прогнозирования его состояния на 5,1% при пороговом значении потери точности до 12%, и отличающийся возможностью обработки гетерогенных данных.

- Разработан алгоритм автоматизированного поиска аномальных значений параметров промышленных манипуляторов, обеспечивающий гибкость процесса мониторинга состояния и отличающийся от штатных программных решений учётом диапазона допустимых значений.

Область исследования соответствует паспорту специальности 2.3.3 «Автоматизация и управление технологическими процессами и производствами» по пунктам:

«6» Научные основы и методы построения интеллектуальных систем управления технологическими процессами и производствами.

«12» Методы создания специального математического и программного обеспечения, пакетов прикладных программ и типовых модулей функциональных и обеспечивающих подсистем АСУТП, АСУП, АСТПП и др., включая управление исполнительными механизмами в реальном времени.

Теоретическая значимость состоит в разработанном комплексе алгоритмов управления историческими данными о работе промышленных роботов и применения методов интеллектуального анализа данных для управления техническим состоянием оборудования. С целью повышения потенциальной эксплуатационной надежности производственного оборудования применяются методы автоматизации обработки многомерных сенсорных данных на основе машинного обучения и искусственных нейронных сетей. Методика

проектирования программного обеспечения, разработанный комплекс алгоритмов и результаты, содержащиеся в настоящей работе, могут быть использованы для управления моделями анализа данных текущего состояния промышленных роботов и планирования работ по техническому обслуживанию.

Практическая значимость результатов исследования заключается в разработанном программном обеспечении, которое реализует требования и ограничения процесса управления техническим состоянием парка разнородного оборудования на основе анализа данных с использованием методов машинного обучения и искусственных нейронных сетей. Разработаны: программное обеспечение «Диагностика аномальных состояний приводов промышленных манипуляторов» (свидетельство о регистрации программы для ЭВМ № 2020614448 от 09.04.2020 г.); «Модуль адаптивного выбора прогнозной модели работоспособности промышленного манипулятора» (свидетельство о регистрации программы для ЭВМ № 2022617263 от 19.04.2022 г.).

Положения, выносимые на защиту

1) Методика проектирования информационной системы для автоматизации предиктивного технического обслуживания и ремонта промышленных роботов на базе проблемно-ориентированного конвейера обработки данных, обеспечивающего повышение экономической эффективности на 5,1%.

2) Алгоритм автоматизированного подбора и оценки эффективности модели машинного обучения, обеспечивающий сокращение временных ресурсов на использование модели анализа данных оценки состояния промышленного робота до 18%.

3) Алгоритм вычисления меры сходства наборов исторических данных о работе промышленного робота обеспечивающий подбор модели анализа данных с потерей точности прогнозирования не более 12%.

4) Алгоритм автоматизированного поиска аномальных значений параметров промышленных манипуляторов, обеспечивающий гибкость процесса

мониторинга состояния за счет формирования диапазона допустимых значений параметров работы на каждой операции рабочей программы.

Апробация и внедрение результатов работы

- Результаты диссертационного исследования использованы при выполнении проекта СТАРТ № 3173ГС1/48576 «Разработка и испытания прототипа программного обеспечения для интеллектуальной диагностики неисправности приводов промышленных манипуляторов», 2019-2020 гг.

- Результаты диссертационного исследования использованы при выполнении проекта РФФИ № 20-37-90113 «Разработка методики адаптивного выбора модели машинного обучения для анализа данных промышленного манипулятора», 2020-2022 гг.

- Результаты диссертационного исследования использованы при выполнении проекта УМНИК № 692ГУЦЭС8-D3/63782 «Разработка алгоритма адаптивного выбора модели машинного обучения для анализа данных промышленного манипулятора с целью оптимизации временных ресурсов», 2020-2022 гг.

- Результаты диссертации использованы в ФГАОУ ВО «Томский государственный университет систем управления и радиоэлектроники» при выполнении государственного задания Министерства науки и высшего образования РФ, проект FEWM-2023-0013 «Гибридная методология построения цифровых моделей социально-экономических и технических систем со структурной и параметрической неопределенностью».

- Результаты диссертационной работы внедрены и используются в ООО «Инспайр-технологии» (г. Томск) – в составе программного обеспечения мониторинга работоспособности и оценки текущего состояния промышленного манипулятора. Акты о внедрении приложены к диссертационной работе.

Основные положения и результаты диссертационной работы доложены и обсуждены на следующих научно-технических конференциях:

- XX Международная научно-практическая конференция студентов, аспирантов и молодых ученых, Томск, 20–22 марта 2023 года. – Томск: Национальный исследовательский Томский политехнический университет.
- XXVII Международная научно-техническая конференция студентов, аспирантов и молодых учёных «Научная сессия ТУСУР – 2022», г. Томск, 18-20 мая 2022 г.
- XVIII Международная научно-практическая конференция старшеклассников, студентов и аспирантов «Молодёжь и наука» 27 мая 2022 года, УрФу, г. Нижний Тагил.
- Международная научно-практическая конференция «Электронные средства и системы управления» (МНПК ЭСиСУ-2020), г. Томск, 18-20 ноября 2020 г.
- 14th International Forum on Strategic Technology (IFOST 2019) 14th-17th October 2019, Tomsk, Russian Federation.
- Всероссийский инженерный конкурс (ВИК – 2019). Симферополь 16-19 декабря 2019 г.; ФГАОУ ВО «КФУ им. В.И. Вернадского».
- XX Всероссийская конференция молодых учёных по математическому моделированию и информационным технологиям. 28 Октября-1 Ноября 2019 г. - Новосибирск: ИВТ СО РАН.
- Форум новых решений UNOVUS-2019. Г. Томск 15-17 мая 2019 г.

Обоснованность и достоверность. Результаты диссертационной работы получены при корректном и обоснованном применении методов оценки сходства данных, анализа данных (машинного обучения и искусственных нейронных сетей), имитационного моделирования и подтверждаются проведенными экспериментальными исследованиями на исторических данных о работе парка из разных промышленных роботов, собранных в процессе рабочего цикла не менее 40 часов.

Публикации. По результатам диссертационной работы опубликовано 12 работ, в том числе 3 статьи в изданиях, включенных в перечень ВАК и 1 работа в

зарубежных изданиях, индексируемых в базах научного цитирования Scopus. Получено 2 свидетельства о государственной регистрации программы для ЭВМ.

Личный вклад автора.

1. Постановка цели и задач исследования осуществлялась совместно с научным руководителем А.О. Савельевым.
2. Программная архитектура проблемно-ориентированного конвейера обработки данных для принятия решений по ТОиР спроектирована автором.
3. Алгоритмическое и программное обеспечение для рекомендации и оценки эффективности модели машинного обучения на основе вычисления меры сходства наборов данных разработано автором.
4. Алгоритм автоматизированного поиска аномальных значений параметров промышленных манипуляторов разработан автором.
5. Сбор и предварительная подготовка данных для апробации программного обеспечения осуществлялась С.С. Михалевичем.
6. Апробация программного модуля рекомендации и оценки эффективности модели машинного обучения произведена совместно с А.Ю. Чепкасовым и А.С. Писанкиным.
7. Разработка имитационной модели для оценки экономического эффекта внедрения разработанного программного обеспечения выполнена автором.

Структура и объем работы. Научная работа состоит из введения, четырех глав, заключения, списка условных обозначений, а также библиографического списка из 167 наименований и 12 приложений. Полный объем диссертации – 185 страниц текста с 80 рисунками и 13 таблицами.

1. Актуальные подходы и методы решения задач оценки работоспособности и состояния промышленных роботов

1.1. Цифровизация производства как элемент автоматизированного управления технологическими процессами

В связи с интенсивным развитием информационных технологий в рамках 4-ой промышленной революции, появляется всё больше высокотехнологичных промышленных производств [32]. Переход на новую ступень развития производства ознаменует появление ряда вызовов технологического уклада, что в свою очередь обуславливает спрос на разработку инновационных решений. Одним из ключевых вызовов является разработка, оснащение и внедрение технологий искусственного интеллекта в промышленное оборудование и производственные процессы [33]. Актуальной остается задача по обеспечению безотказного функционирования промышленного оборудования при минимизации затрат на его содержание и техническое обслуживание.

Процессы технического обслуживания в различных отраслях промышленности рассматривались такими российскими и зарубежными авторами как: А.А. Мисловидова, И.О. Бельский, К.С. Ахмедов, А.А. Сычугов, И.В. Сорокина, А.Е. Савочкин и др. Вклад в исследования по разработке и реализации методов интеллектуального анализа данных для процесса технического обслуживания внесли: И.С. Макаш, С.В. Квонг, А.В. Суханов, М.Ш. Омар, И.П. Врублевский, О.Н. Долинина, Е.Е. Дышкант, А.В. Кизим, L. Zhang, C.L. Wang, E. Sezer, D. Romero, F. Guedea, R. Chianese, D.K. Saha, S. Ahmed и др. Исследованию и разработке методов прогнозирования работоспособности, ранних отказов и остаточного ресурса промышленного оборудования уделяли внимание следующие ученые: P. Singh, S. Agrawal, J. Dsouza, S. Velan, R. Ashmore, L. Zhang, R. Calinescu, L.P. Silvestrin, G. Ortiz и др. Вызовы и проблемы использования и управлением моделями машинного обучения и большими данными исследовались: S. Schelter, F. Biessmann, T. Januschowski, F. McSherry, A. Kumar, R. J Hyndman и др. Многие

современные отечественные (ООО «Цифра», Ростех, Северсталь, МАК «Вымпел» и др.) и международные (SAP, AT&T, Siemens, General Electric, Microsoft, KUKA, ABB, Fanuc и др.) компании так же развивают направления по предиктивному техническому обслуживанию и принятию решений на основе данных. Основным и наиболее часто используемым прикладным инструментарием специалистов и исследователей, работающих в сфере предиктивной аналитики, являются:

- Языки программирования: R, Python (в том числе: Pandas, NumPy, SciKil-Learn, SciPy, PySurvival, Keras, PyTorch, TensorFlow, Statistics и др.) [1-4];
- Среды управления вычислениями: Anaconda, Google Colab, Matlab, R Software Environment, Apache Mahout, GNU Octave и др. [5-7].
- Программные пакеты анализа данных: Excel, Alteryx APA, TIBCO Spotfire, Orange Data Mining, Knime Analytics, Weka Data Mining, RapidMiner, IBM SPSS Modeler, Watson Analytics, SAS Enterprise Miner, MLFlow, Oracle Big Data Preparation, Sisense и др. [8, 9].
- Средства визуализации: Python (включая программные библиотеки: matplotlib, seaborn, plotly, bokeh), FusionCharts, Power BI, Tableau, Chart Studio, Qlik Sense, DataLens и др. [10].
- Системы накопления и хранения данных: Hadoop Distributed File System (HDFS), YTsaurus, Enterprise Resource Planning (ERP), Hadoop YARN, Customer relationship management (CRM), Human Resource Management Systems (HRMS), Extract-Transform-Load (ETL), Data Warehouse (DWH) и др. [11, 12].

Наиболее часто используемыми методами для анализа данных в рамках предиктивной аналитики являются:

- Линейная и логистическая регрессия [13, 14].
- Деревья решений (Decision Trees) и случайный лес (Random Forest) [15, 16].
- Авторегрессия и авторегрессионная интегрированная скользящая средняя (ARIMA) [17-19].

- Наивный Байес, Метод опорных векторов (SVM), k-ближайших соседей, DBSCAN [20-23].
- Цепи Маркова, Скрытые Цепи Маркова [24-26].
- Нейронные сети (Auto-encoder, Recurrent Neural Networks, Long short-term memory и др.) [27-31].

Интеллектуализация производственных процессов затрагивает следующие аспекты:

- Автоматизация производственной деятельности на базовом уровне – наличие информационных систем для сбора данных, обратной связи и управления конкретной производственной единицей оборудования.
- Управление энергией и ресурсами – внедрение систем мониторинга, анализа и оптимизации использования энергии, производственных ресурсов и оборотных средств.
- «Интеллектуальное» производство – наличие системы сбора и анализа данных, содержащих наиболее комплексное описание бизнес-процессов промышленного производства, служащей для поддержки принятия стратегических решений при планировании и реинжиниринга процессов производства.
- «Интеллектуальное» оборудование – основные средства, оснащенные системами для мониторинга, сбора и анализа данных, а также программными модулями для формирования целостной производственной линии на основе стандартизованных программных интерфейсов.

Для создания «интеллектуального» промышленного производства необходимо стремиться к соблюдению следующих требований:

- Требования на физическом уровне – наличие у основных средств систем для сбора данных (в том числе «умных датчиков» [35]) и высокоскоростных интерфейсов передачи данных.
- Требования на уровне передачи данных – поддержка современных протоколов передачи данных, позволяющих настраивать оптимальные по структуре и масштабируемые локальные промышленные сети.

- Требования на прикладном уровне – обеспечение сбора данных в системы хранения (в том числе с использованием облачных технологий), реализующих интерфейсы взаимодействия с системами интеллектуального анализа данных и поддержки принятия решений.

В общем виде многоуровневая система современного цифрового производства представлена на рисунке 1. Интеллектуализация процессов производства приводит к следующим эффектам:

- Увеличение производительности – цифровизация и ИИ позволяют автоматизировать рутинные задачи, оптимизировать рабочие процессы и улучшить эффективность производства, что приводит к увеличению общей производительности и сокращению времени, затрачиваемого на выполнение операций.

- Снижение затрат – оптимизация использования ресурсов, управление складскими запасами и логистикой, снижение затрат на производство.

- Улучшение качества продукции – применение методов ИИ в процессе производства позволяет более точно контролировать и отслеживать каждый шаг производственного процесса.

- Прогнозирование – в процессе производства и хозяйственной деятельности генерируется большой объем исторических данных, применение методов анализа данных к которому позволяет извлекать ценные знания. Результаты анализа позволяют предсказывать спрос на продукцию, оптимизировать планирование производства и прогнозировать неисправности оборудования, что способствует более эффективному управлению и принятию решений.

- Повышение уровня надежности – цифровизация и внедрение ИИ позволяет создавать системы мониторинга и контроля, которые способствуют прогнозированию и предотвращению аварий, что в свою очередь минимизирует риски для сотрудников и незапланированных простоев производства.

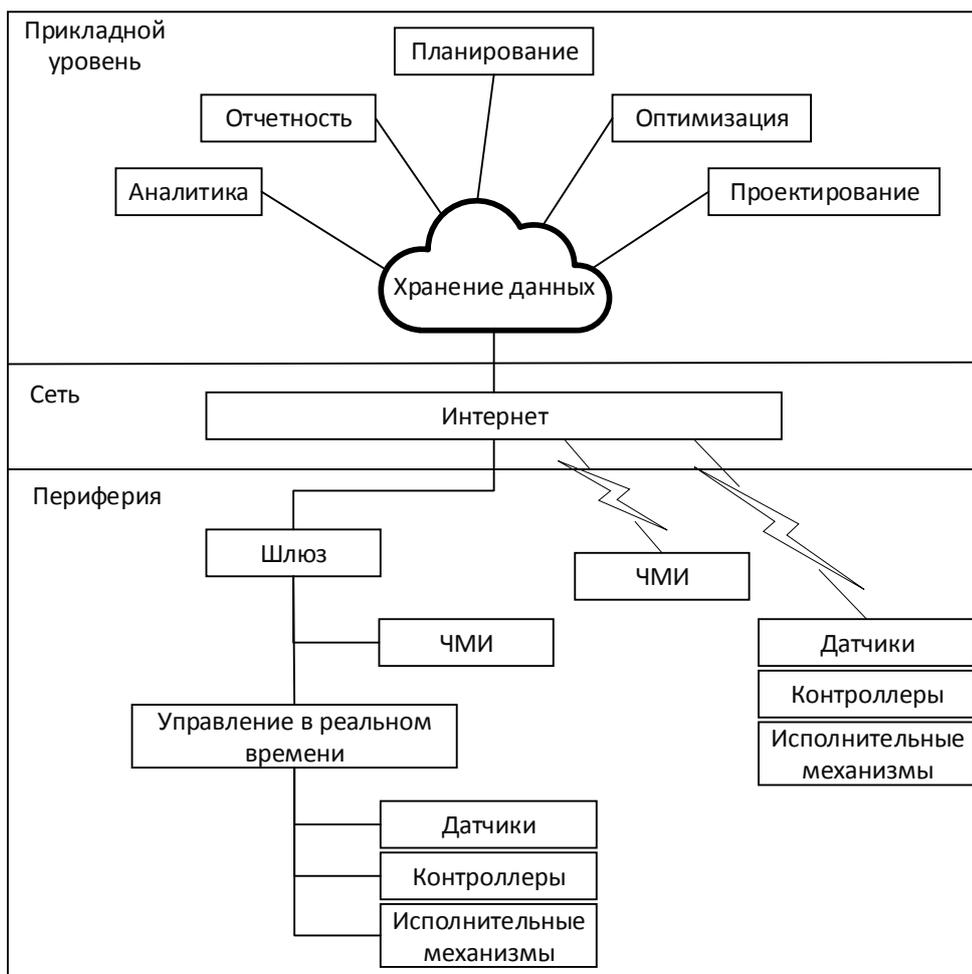


Рисунок 1 – Представление цифрового производства

Цифровизация производства позволяет реализовать модель кибер-физической системы (CPS – Cyber Physical System или CPPS – Cyber Physical Production System) [36, 37]. Такого рода системы объединяют в единую сеть информационные и физические компоненты, функционирующие с использованием стандартизованных протоколов передачи данных и алгоритмов взаимодействия настройки, обратной связи, управления и адаптации к изменяющимся правилам и внешним условиям. Выделяется пять уровней построения кибер-физической производственной системы:

1) Соединение – коллекционирование данных, генерируемых подключенными в сеть оборудованием, системами контроля и мониторинга (ERP, CRM, HRMS и др.).

2) Преобразование – подготовка данных (ETL, DWH и др.) для последующего анализа программными пакетами.

3) Кибер-уровень – использования алгоритмов анализа больших данных о работе парка оборудования, включая рекомендательные системы по оптимизации использования отдельных активов; обеспечение отображения парка в цифровом виде (Digital Twin).

4) Знания – автоматизированный онлайн-мониторинг и диагностика потенциальных отказов оборудования на ранней стадии. Основываясь на адаптивном обучении с использованием исторических оценок работоспособности, подсистема анализа данных реализует специальные алгоритмы для прогнозирования и оценки времени до отказа,

5) Конфигурация – самодиагностика оборудования для обнаружения отказов на ранних этапах с предоставлением отчетов о работоспособности на операционный уровень. Организация обратной связи и саморегулирования рабочей нагрузки.

1.2. Подходы и методы интеллектуализации процессов управления технологическими процессами и производствами

Одновременно с системами сбора данных развиваются инструменты их обработки и анализа с целью извлечения полезных знаний, позволяющих выявлять возможности по оптимизации производственных издержек. Развертывание усовершенствованной сетевой коммуникационной инфраструктуры, методов высокопроизводительного анализа данных (High-Performance Data Analytics) и возможностей высокопроизводительных вычислений (High Power Computing) в облаке и на периферии позволяет перейти к использованию подхода вычислительного интеллекта для детектирования «аномалий», прогнозирования отказов и создания цифровой «гиперсвязи» [39].

В частном случае это достигается за счет управления процессом планирования технического обслуживания основных средств [40]. Принятие

решений при техническом обслуживании основывается на извлечении и получении новых знаний из наборов данных. Новое знание в обслуживании должно быть сосредоточено на двух аспектах:

- Какую информацию возможно извлечь в настоящих условиях?
- Какая информация должна быть известна для полноты и достаточности формирования решений?

Сбор данных может производиться в разных режимах (в зависимости от субъекта управления):

- 1) Инспектирование человеком – процесс подразумевает присутствие человека с различной квалификацией.
- 2) Удаленная проверка – выполнение с использованием удаленных платформ через подключение к сети.
- 3) Проверка механизмом – выполнение с использованием другого оборудования, а именно: специализированное средство контроля, автономный сенсорный робот или программная система контроля.

Способы технического вмешательства для обслуживания зависят от конструкции и функций, обеспечиваемых самой системой, а также от ее размера, стоимости и условий технической безопасности:

- 1) Перепрограммирование: оборудование перезагружается (либо происходит сброс до заводских настроек) или параметризуется заново для обеспечения необходимого рабочего функционала.
- 2) Замена деталей: неисправные или изношенные детали системы заменяются новыми с целью восстановления целевых функций.
- 3) Полная замена: вся система заменяется в следствии потери функциональности без возможности восстановления по практическим или экономическим причинам.

В целях максимальной экономии учитываются все затраты, связанные с вышеуказанными способами проверки. С точки зрения управления техническим обслуживанием затраты можно разделить на постоянные и переменные в зависимости от воздействия на систему.

Концепция Maintenance Analytics (МА) объединяет подходы по обнаружению знаний для технического обслуживания с упором на большие данные и аналитику [135]. МА фокусируется на синтезе новых знаний в обслуживании для прогнозирования потенциальных неисправностей и определения оставшегося срока службы оборудования [41]. Модели, реализующие оценку остаточного полезного ресурса оборудования, не только прогнозируют срок службы, но и обеспечивают доверительную границу предсказания. Техническое обслуживание на основе состояния – это стратегия принятий решений по обслуживанию оборудования на основе данных о его фактическом состоянии [42-45]. Для описания стратегий технического обслуживания выделяют следующую классификацию:

- Реактивное обслуживание (Reactive maintenance) – обслуживание оборудования после случившегося отказа.
- Превентивное обслуживание (Preventive maintenance) – обслуживание согласно утвержденному регламенту [46].
- Обслуживание на основе состояния (Condition-based maintenance) – обслуживание, основанное на непрерывном мониторинге состояния и работоспособности оборудования и отдельных узлов.
- Обслуживание на основе статистики (Statistic-based maintenance) – обслуживание, основанное на непрерывном мониторинге состояния и использовании инструментов анализа данных, использующих методы прогнозирования состояний: машинное обучение, статистические выводы и инженерные подходы.

Техническое обслуживание на основе состояния базируется на мониторинге оценки текущего состояния оборудования. Данный подход включает в себя отслеживание состояния актива с использованием определенных показателей эффективности. Перечень методов, наиболее часто применяемых в мониторинге состояния:

- Виброакустический анализ – данный тип мониторинга состояния выявляет потенциальную неисправность, обнаруживая изменения в характеристике вибрации узлов, корпуса и прочих составляющих [47-50].
- Термический анализ – мониторинг нагрева агрегатных узлов, деталей и систем для предотвращения износа, деформации, поломки или возгорания [51-53].
- Ультразвуковой анализ – мониторинг износа с использованием специализированного оборудования, позволяющего фиксировать ультразвук, передающийся по воздуху или по структуре материалов, и преобразовывать его в слышимые сигналы, которые воспринимает специалист в наушниках и видит на индикаторной панели в виде графиков [54-57].
- Акустический анализ – использование датчиков и микрофонов для обнаружения звуков, указывающих на нештатную работу оборудования. Акустический анализ имеет дополнительное преимущество, заключающееся в возможности выявления утечек газа, жидкости или вакуума [58, 59].
- Анализ электрического тока – мониторинг отклонений значений тока, мощности, напряжения от штатных при работе приводов, двигателей, контроллеров [62, 63]. Целью контроля технического состояния является оценка целостности агрегата и раннее предупреждение о возможных неисправностях.

В статье [64] приведена работа по систематизации документов, участвующих в процессе управления состоянием и обслуживанием оборудования (Рисунок 2 – Процесс управления техническим обслуживанием).

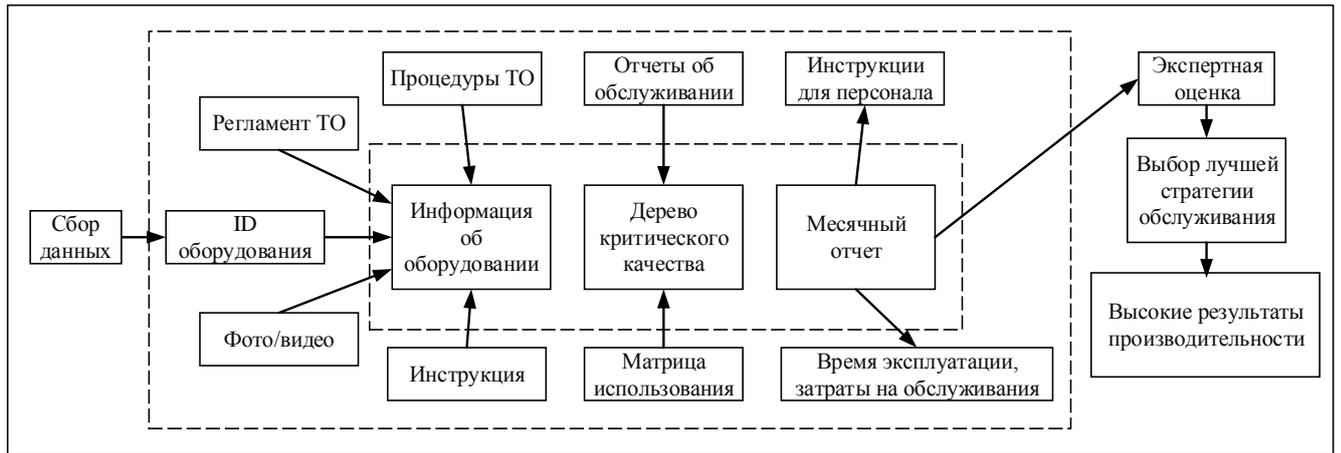


Рисунок 2 – Процесс управления техническим обслуживанием

Имеющийся опыт показывает, что проблема управления промышленным оборудованием является комплексной, включающей технической, технологической, управленческой, экономической и организационный аспекты, требующие разработки методов и алгоритмов автоматизации процессов мониторинга и прогнозирования технического состояния контролируемого оборудования в реальном времени [65]. Непрерывный процесс получения данных и своевременный анализ позволяет реализовывать комплексные меры для минимизации устранения возможных рисков:

- 1) Выявление проблем.
- 2) Оценка основных областей риска.
- 3) Измерение вероятности реализации риска.
- 4) Ранжирование рисков.
- 5) Определение желаемых результатов.
- 6) Разработка вариантов работы с рисками.
- 7) Выбор стратегии.
- 8) Реализация стратегии.
- 9) Мониторинг, оценка и коррекция.

С развитием методов интеллектуального анализа (Data Mining) появляется больше технологических решений для создания «умной фабрики» или «умного» производства [66]. Реализация такого рода систем позволяет перейти к стратегии управления на основе данных, применяемой для прогнозирования и планирования

технического обслуживания основных средств [67, 68]. Для оценки достоверности результатов работы и улучшения процесса предиктивного технического обслуживания используется подход Process Mining [69]. Данный подход предложен в качестве описательного метода для обнаружения закономерностей и детального понимания бизнес-процессов. Результаты применения Process Mining используются для оптимизации бизнес-процессов за счет выявления «узких мест», систематического нарушения нормативных требований и выявления отклонений от заданного процесса. В связке с методами интеллектуального анализа данных данный подход позволяет реализовать:

1) Кластеризацию результатов интеллектуального анализа процессов – использование методов машинного обучения для группировки схожих событий и кейсов в кластеры.

2) Обнаружение «аномалий» – использование методов машинного обучения для обнаружения «аномальных» значений (выбросов, сдвигов, изменения распределения значений, отклонений от рабочей программы).

3) Сходство результатов анализа – использование методов машинного обучения для нахождения схожих кейсов (может согласоваться с подходом Case-Based Reasoning).

Внедрение методов анализа данных в рамках предиктивного обслуживания влечет за собой следующие преимущества:

- Прогнозирование отказов – применение методов анализа данных и искусственного интеллекта позволяет анализировать большой объем данных о состоянии оборудования и процессе работы.

- Оптимизация планов технического обслуживания – позволяет оптимизировать расписание технического обслуживания и разрабатывать наиболее эффективные стратегии обслуживания на основе данных, параметров работы оборудования и его состояния.

- Экономическая эффективность – выявление проблем на ранней стадии позволяет предотвратить серьезные поломки и связанные с ними высокие расходы на ремонт, замену оборудования и незапланированные простои.

Проведенные исследования в области мониторинга работы электродвигателей широко охватывают основные проблемы и неисправности, возникающие в результате эксплуатации оборудования [70, 71]. В исследованиях приводится классификация наиболее часто встречающихся неисправностей, связанных как с внешними причинами (условия производства, качество обслуживания и мониторинга), так и с внутренними (связанные с неисправностями в строении оборудования). В исследованиях делается акцент на потребности в дальнейшем развитии систем сбора и анализа данных для более эффективного прогнозирования отказов оборудования.

Под неисправностями в работе часто могут подразумеваться аномальные значения, получаемые в процессе мониторинга оборудования. Методы анализа данных позволяют расширять возможности применения выявления аномалий не только в контексте неисправности или раннего отказа, но и для оценки качества выполняемого производственного процесса [157]. В исследовании [158] предлагается использование методов классификации параметров мониторинга для разметки «аномальных» и «штатных» ситуаций при агрегировании пакетов данных, получаемых с промышленных роботов. Предлагаемый подход позволяет гибко отслеживать состояние безопасной работы манипулятора (то есть в пределах допустимой нагрузки) и оперативно сигнализировать о возникших аномалиях, посредством интеграции в информационную инфраструктуру производства. В работе [159] описывается разработанная программная система для выявления аномальных значений при работе промышленных манипуляторов с учетом специфики производственного процесса (т.е. заданной оператором программы). Предложенный подход позволяет с высокой точностью производить контроль движений манипулятора по заданным программой траекторий.

Одним из подходов диагностики отказов является «Рассуждение на основе прецедентов» (Case-Based Reasoning). Данные для анализа собираются и используются для классификации сбоев оборудования, происходивших в процессе мониторинга [72-74]. Использование данного подхода предполагает решение новых проблем (кейсов) с использованием готовых или похожих решений

(сформированных планов диагностики неисправностей на основе произведенной классификации сбоев). Применительно к промышленным роботам данный метод применяется следующим образом:

- 1) Производится описание и стандартизация возможных отказов (сбоев, поломок).
- 2) Характеристики возможных отказов выражаются в форме признакового описания (feature vector).
- 3) При обнаружении отказа с близким признаковым описанием в пределах порогового значения применяется соответствующее существующее решение для диагностики и разрешения возникшей проблемы с оборудованием.
- 4) В ином случае, план диагностики подвергается модификации в соответствии с наиболее похожими случаями отказов для формирования нового плана диагностики.

Метод рассуждения на основе прецедентов представлен на рисунке 3.

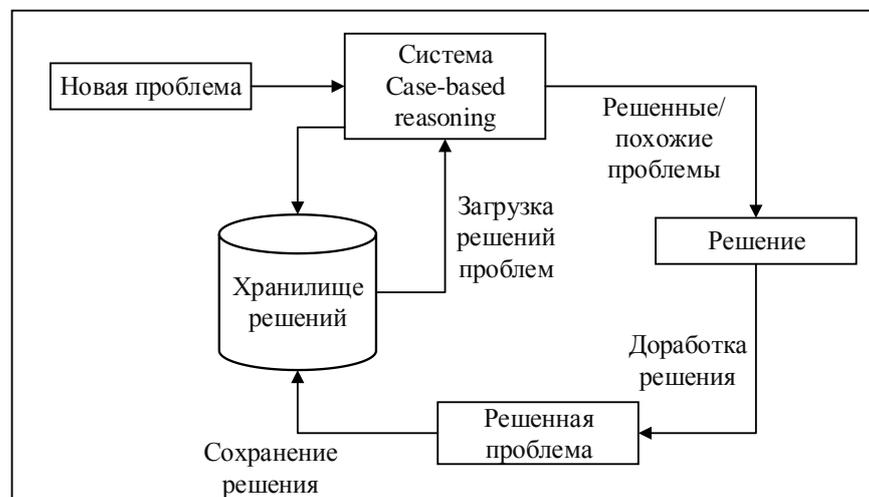


Рисунок 3 – Метод рассуждения на основе прецедентов (кейсов)

С развитием систем сбора данных (датчиков, протоколов передачи и агрегирования) становится возможным производить комплексный мониторинг и анализ исторических данных оборудования. Актуальной тенденцией технического обслуживания является прогнозный подход, примером которого являются методы прогнозирования и управления здоровьем (Prognosis and Health Management),

технического обслуживания на основе состояния (Condition-based Maintenance) и остаточный срок службы (RUL – Remaining Useful Life) [75, 76]. Точное прогнозирование остаточного срока службы необходимо для принятия эффективных решений для прогнозного обслуживания и/или для технического обслуживания на основе состояния (рисунок 4).

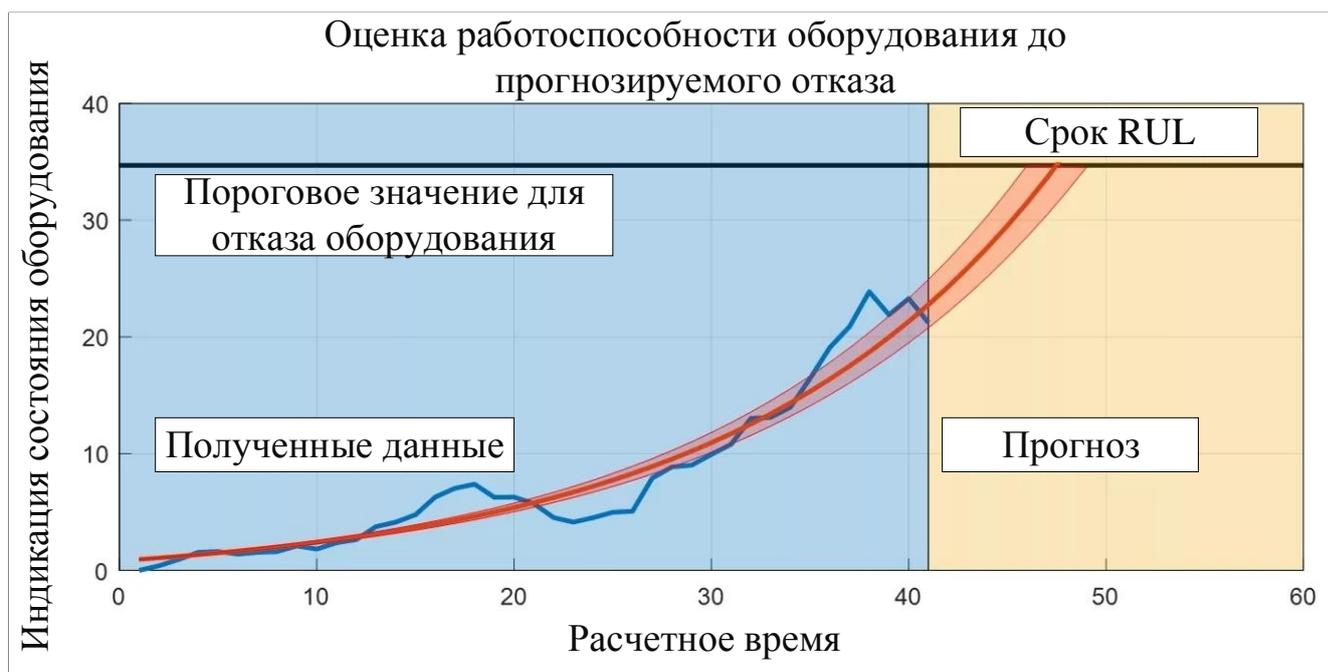


Рисунок 4 – Концептуальное представление прогнозирования остаточного срока работоспособности оборудования с помощью подхода Remaining Useful Life

RUL определяется как условно случайная величина:

$$T - t | T > t, X(t), \text{ где} \quad (2)$$

где, T – время до поломки, t – текущее время работы, $X(t)$ – исторический «профиль» работы оборудования (данные мониторинга состояния). Увеличение точности RUL достигается за счет применение методов статистики, машинного обучения, скрытых марковских моделей, главных компонент и искусственных нейронных сетей [77]. RUL – прогнозируемый период времени, в течение которого оборудование может работать, прежде чем ей потребуется ремонт или замена (рисунок 5). Методы, используемые для оценки RUL, зависят от:

- Данных о сроке службы оборудования и узловых агрегатов.
- Исторических данных, описывающие работу до отказа.
- Пороговых значениях параметров работы оборудования, свидетельствующих о отказе.

Данный подход, как и большинство методов анализа, зависит от качества собираемых данных – проблемы возникают, когда качество входных данных недостаточно высоко или данные содержат выбросы, пропуски или ошибки. Это может привести к неправильным выводам и плохим прогнозам. Оставшийся срок полезного использования компонентов и систем определяется посредством оценки состояния оборудования и используется для расчета ожидаемых краткосрочных и долгосрочных капитальных затрат, необходимых для содержания основных средств. В работе [78] приведен систематизированный обзор подходов RUL на примере оценки оставшегося полезного срока службы режущих инструментов в процессе фрезерования. В работе описаны различные методы мониторинга, методы извлечения признаков, модели принятия решений и доступные датчики, используемые в data-driven моделях.

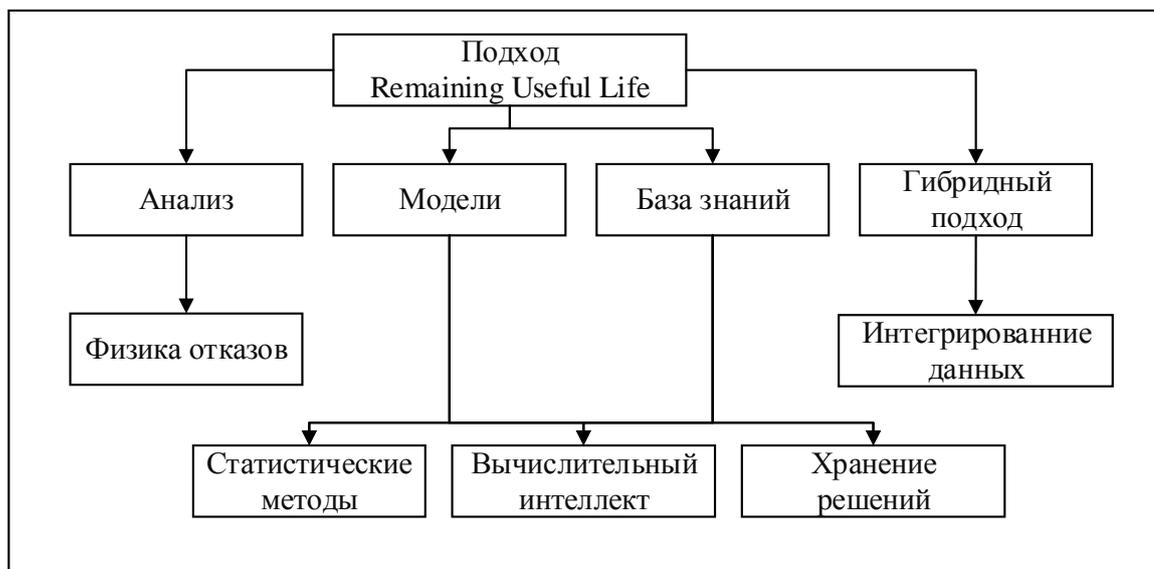


Рисунок 5 – Методы, используемые для оценки остаточного срока службы оборудования

Существует несколько методов прогнозирования, используемых для определения RUL подсистем или компонентов:

- Методы на основе использования моделей – использование статистического подхода и вычислительного интеллекта применительно к конфигурации системы, характера использования и исторических данных.
- Методы анализа для прогнозирования физического отказа компонентов оборудования. Для анализа создается цифровая физическая модель объекта с учетом характеристик и условий эксплуатации (параметрические данные и методы покрытия материала, механика повреждений, правила линейных повреждений, нелинейные кривые повреждений и двухэтапная линейаризация).
- Методы на основе знаний – применение комбинации из вычислительного интеллекта и используемых опытных решений.
- Гибридная модель использует параметрические и непараметрические данные для вычисления оценок RUL с целью повышения точности прогнозирования. Результатом работы модели является значение порога RUL индивидуально для каждой единицы оборудования.

Процесс принятия решений при проактивном обслуживании нуждается в большом объеме собранных данных. Достоверно определенный тип отказа влечет за собой принятие правильных мер по техническому обслуживанию. Классификация отказов выполняется путем внедрения знаний, полученных из результатов анализа видов и последствий отказов (Failure Modes and Effects Analysis, FMEA) и результатов физики отказов (Physics-of-Failure, PoF или Fault-Based, FB) [78]. Подход FB основан на рабочем процессе с использованием методов ML для интеллектуального анализа данных и дополнительном этапе планирования технического обслуживания. На этапе сбора данных выполняется множественное тестирование для создания набора данных, в достаточной мере описывающего множество возможных аномальных состояний и неисправностей системы. На этапе установки цели используются априорные знания об ошибках для определения целевого вектора, предполагая значения в наборе из n элементов, где n – количество типов ошибок. На финальном этапе управления техническим обслуживанием

производится перекрестная связь списка неисправностей, полученного в результате классификации неисправностей, с анализом затрат, с целью планирования правильных операций по техническому обслуживанию в рамках регламента.

Подход, основанный на плановом обслуживании, аналогичен модели на рисунке 6, но в приоритет ставится регламент обслуживания и конечный набор действий, который предполагается как минимально необходимый для работы с выявленным типом неисправности. Данное предположение проверяется в соответствии с требованиями анализа расходов, проводимого субъектом управления, утверждающего перечень работ по техническому обслуживанию.

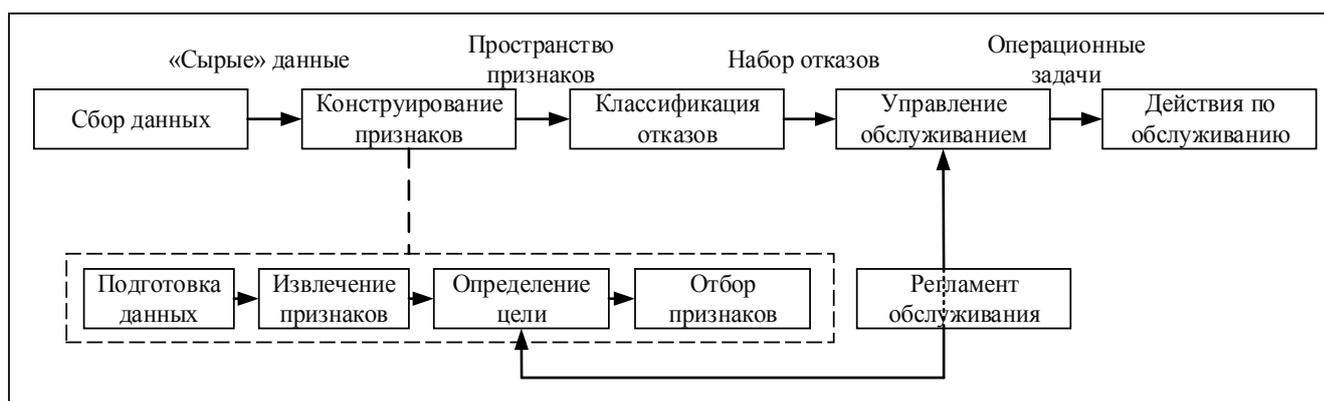


Рисунок 6 – Процесс управления техническим обслуживанием на основе анализа данных

В исследованиях [138-140] рассматривается технология цифрового двойника, включая архитектуру «экосистемы» и применение данной технологии в рамках системы прогнозного технического обслуживания (рисунок 7). В исследованиях отмечаются такие проблемы как: слияние данных (несмотря на технические сложности в интеграции разнородных цифровых и аналоговых данных с оборудования, остаются неисследованными вопросы по наиболее адаптивной и оптимальной интеграции цифровых данных и результатов принятия решений живым человеком), непрерывного обучения и адаптация моделей анализа данных под изменяющиеся условия или ограничения при внесении дополнительных

изменений в рабочий процесс. В качестве развития подхода прогнозного технического обслуживания предлагается подход предписывающего обслуживания (prescriptive maintenance), что подразумевает автоматизированную разработку рекомендаций по планированию расписания и операционных действий по обслуживанию.



Рисунок 7 – Концептуальная архитектура рекомендательной системы для принятия решений на основе состояния оборудования

В исследовании [141] представлена методология, позволяющая внедрять цифровые двойники промышленных роботов и сложных современных производственных станков. Реализация и внедрение DT может обеспечить ряд преимуществ, например, в сочетании с кривыми ухудшения критических показателей (DCA) работоспособности позволяет прогнозировать RUL 6-осевых

роботов (рисунок 8). В работе приводится анализ общей методологии для решения таких задач, как:

1. Определение параметров физической модели для ДТ.
2. Выбор параметров для настройки разрабатываемого программного модуля двойника и частоты его обновления для учета физических изменений в цифровой модели.
3. Расчет и моделирование новых значений параметров, на основе собранных архивных данных и их автоматизированный перенос на цифровую модель.

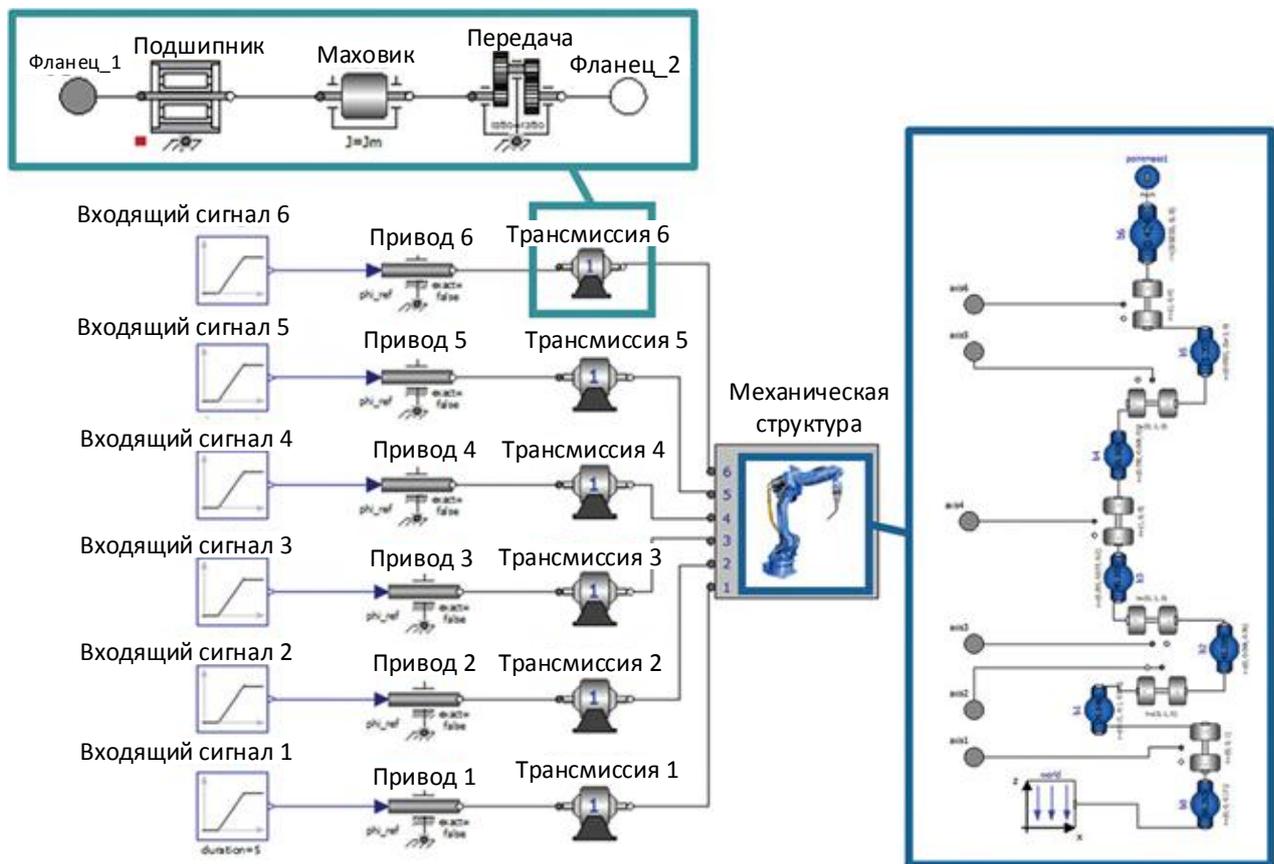


Рисунок 8 – Схематичное изображение составных частей 6-осевого манипулятора

1.3. Подходы и ограничения процессов интеллектуальной оценки состояния промышленных роботов

Методы машинного обучения все чаще применяются в разных областях управления. В исследованиях [80, 81] представляется обзор на традиционные и современные методы обслуживания производственного оборудования, используемые для прогнозирования работоспособности и повышения производительности производства. С точки зрения машинного обучения существует несколько проблем профилактического обслуживания производственных линий. Во-первых, сложности в получении достаточного объема параметров описания данных о неисправностях оборудования и маркировка случаев отказа в наборе данных. Во-вторых, существует огромное количество технологических данных, генерируемого на производственных линиях и роботизированных ячейках, и для обработки больших данных требуется специальная информационная инфраструктура, экспертные знания и специализированное программное обеспечение, реализующее методы интеллектуального анализа данных.

В общем виде структурно-функциональная схема автоматизированной системы управления процессом технического обслуживания и ремонта промышленного робота представлена на рисунке 9.



Рисунок 9 – Структурно-функциональная схема автоматизированной системы управления техническим обслуживанием промышленного робота

Объектом управления является промышленный робот, представляющий из себя кибер-физическую систему с интерфейсами управления и устройствами сбора данных. В процессе эксплуатации оборудования производится генерация данных Y с помощью устройств сбора данных (датчики, сервоприволы, контроллеры) и системы мониторинга, которая агрегирует и интерпретирует входные данные для формирования выходных пакетов с накопленными данными Y^* . Выходные данные из систем мониторинга являются объектом изучения для прогнозирования остаточного срока службы оборудования, его составляющих и прогнозирования потенциальных отказов в рамках процесса предиктивного технического обслуживания и ремонта. Методы статистического и интеллектуального анализа являются неотъемлемой частью предиктивного анализа. Данные, полученные в результате анализа, используются для процесса принятия решений при техническом обслуживании и ремонта с экономической оценкой возможных альтернатив (инспекция, настройка, ремонт или замена оборудования) в зависимости от: стоимости операции, наличия комплектующих, целесообразности проведения тех или иных работ. На основе принятого решения формируется план,

включающий в себя предписания по техническому обслуживанию и ремонту. Руководствуясь предписаниями по техническому обслуживанию, исполнительными устройствами формируются управляющие воздействия U , представляющие из себя плановые или корректирующие операции. Модель системы устройства управления предиктивным техническим обслуживанием промышленного робота определена следующей формой:

$$\begin{cases} \dot{X} = AX + BU \\ Y = CX + DU, \end{cases} \quad (1)$$

$$U = f_{yy}(Y^*), \quad Y^* = \begin{bmatrix} Y_1^* \\ Y_2^* \\ \dots \\ Y_n^* \end{bmatrix}, \quad \text{где} \quad (2)$$

X – переменная состояния объекта управления;

\dot{X} – производная управляемой величины от задающего воздействия;

Y – сигналы, полученные из устройств сбора данных объекта управления;

Y^* – накопленные данные, получаемые с использованием системы мониторинга;

U – управляющее воздействие исполнительного устройства на объект управления;

A, B, C, D – матрицы: системы, управления, наблюдения, связи;

f_{yy} – функция управления техническим обслуживанием и ремонтом на основе методов анализа данных.

Матрица системы A , элементы которой определяются структурной схемой системы и значениями ее параметров, характеризует динамические свойства системы, ее свободное движение. Матрица управления B характеризует влияние внешних воздействий на переменные состояния системы, т.е. определяет чувствительность системы к внешним воздействиям (задающим и возмущающим). Матрица наблюдения C характеризует связь выходной величины системы с вектором состояния. Обычно не все составляющие вектора состояния являются

наблюдаемыми сигналами, т.е. могут быть измерены с помощью каких-либо датчиков, в то время как выходной сигнал всегда наблюдаем. Матрица связи D устанавливает связь выходной величины системы с внешним воздействием.

Выбор предписаний для управляющего воздействия U определяется результатами обработки данных методами анализа данных (статистическими, машинного обучения и ИНС).

Одним из основных понятий машинного обучения является обучающая выборка:

$$X = (x_i, y_i)_{i=1}^l, \text{ где} \quad (3)$$

x_i – признаковое описание объекта управления X ,

y_i – целевой параметр на объекте,

l – размер выборки.

Признаковое описание представляет собой набор физических данных, собираемых системами мониторинга при выполнении рабочего цикла. Целевым параметром является классификация состояний объекта управления. Прогнозирование состояния основывается на применении модели $a(x)$, представляющей собой функцию из пространства объектов управления X (парка оборудования) и пространство ответов целевых показателей состояния Y . Эксплуатация модели предиктивной аналитики происходит в три основных этапа: получение данных в режиме реального времени, конструирование и использование предсказательной модели, интерпретация результатов использования модели в понятной для субъекта управления форме. Эффективность процесса планирования мероприятий зависит не только от своевременно полученных результатов обработки данных, но и корректно сформированных правил использования данных для обоснования выбора оптимального решения при оценке альтернатив.

В работе [82] рассматривается задача построения предсказательных моделей для решения задач индустриальной предиктивной аналитики. Автор, проанализировав потребности индустриальных приложений, сформулировал ряд

новых математических и алгоритмических задач и разработал соответствующие методы моделирования на основе данных. Предложен подход к планированию экспериментов для разнородных данных, основанный на алгоритмах сглаживания [83]. Авторами разработан программный комплекс для построения метамоделей на основе разнородных данных, в том числе большого размера [84].

Глубокое обучение – использование методов искусственных нейронных сетей (таких как: Auto-encoder, Deep Belief Network, Deep Boltzman Machine, Convolutional Neural Networks, Recurrent Neural Networks и др.) для построения моделей прогнозирования отказов и будущего состояния оборудования [85]. Преимуществами данного подхода выделены:

- Работа на большом объеме исторических данных – чем больше данных, тем более точные получаются модели.
- Выделение признаков высокого уровня для анализа (классификация, распознавание, сегментация и др.).
- Работа модели по принципу end-to-end – решение проблемы с непосредственной обратной связью для конечного пользователя.

В исследовании [86] проведено сравнение методов машинного и глубокого обучения с целью оценки эффективности применения для профилактического подхода к обслуживанию гидравлической системы. Архитектура долгой краткосрочной памяти (LSTM – тип рекуррентной нейронной сети, способный обучаться долгосрочным зависимостям) применяется для словесного перевода и задач моделирования. Применение LSTM для анализа коротких временных последовательностей не приводит к удовлетворительной точности прогнозирования в отличие от временных сверточных сетей (Temporal Convolutional Network).

$$F(t) = (x * df) \cdot (t) = \sum_{i=0}^{k-1} f(i) \cdot x_{t-d*i}, \text{ где} \quad (4)$$

t – временной шаг свертки входного ряда x , свернутого фильтром $f:\{1..k\} \rightarrow \mathbb{R}$. На каждом шаге t производится поэлементное умножение между фильтром и k

предыдущих элементов x . Коэффициент расширения d – это количество шагов по времени, которые пропускаются между каждым из k элементов x , участвующих в свертке до t . Данный метод также эффективен для оценки RUL (71% точности модели для прогноза RUL на авиационных силовых установках) [87].

Одним из эффективных методов прогнозирования временных рядов, используемых для мониторинга потенциальных неисправностей, является ARIMA (авторегрессионная интегрированная скользящая средняя). В работе [88] данный метод использован для анализа неисправностей в режиме реального времени и прогнозирования трендов, приводящих к отказу. В исследованиях [89, 90] данный метод применен совместно с рекуррентной нейронной сетью для анализа факторов, влияющих на прогнозирование количества пользователей программного продукта (метриками оценки моделей являются симметричная средняя абсолютная процентная ошибка и средняя квадратичная ошибка). Метод ARIMA описывается следующей математической моделью:

$$ARIMA(p, d, q): y(t) = c + \sum(\varphi(i) \cdot y(t - i)) + \sum(\theta(i) \cdot \varepsilon(t - i)) + \varepsilon(t), \text{ где (5)}$$

$y(t)$ – значение временного ряда в момент времени t ,

c – константа,

$\varphi(i)$ – коэффициенты авторегрессии,

$\theta(i)$ – коэффициенты скользящего среднего,

$\varepsilon(t)$ – случайная ошибка в момент времени t ,

p, d, q – параметры модели ARIMA, определяющие количество лагов авторегрессии, разности и скользящего среднего соответственно.

Аномальные значения параметров при работе промышленных роботов могут свидетельствовать о потенциальном отказе оборудования. В исследовании [91] приводится обзор и применение следующих методов анализа данных для оценки состояния промышленного манипулятора:

- Частотный анализ временного ряда вибраций привода.

- Смещение среднеквадратичного значения для временного ряда данных тока.
- Метод опорных векторов и ИНС для классификации состояний системы.

В исследовании по разработке системы анализа аномальных данных о работе промышленного робота предлагается следующее решение – подсистема создания и обучения модели, которая описывает состояние робота на основе исторических данных при штатной эксплуатации [92]. Для классификации аномальных значений применяется метод опорных векторов в совокупности с сетью радиально базисных функций. С целью оптимизации ресурсов для обучения модели посредством уменьшения размерности данных используется метод анализа главных компонент [93]. Данный метод аппроксимирует n -размерное пространство данных до n -мерного эллипсоида, полуоси которого являются результирующими главными компонентами. Обучение и настройка моделей анализа данных может занимать значительное время, особенно в случае больших объемов данных или сложных моделей. Применение вышеописанного метода решает данную проблему при возникающей необходимости своевременного обновления и дообучения моделей.

Применение прогнозного моделирования для прогнозирования состояния промышленного робота в работе [75] производится с целью выявления ошибок точности при работе манипулятора. Предлагаемый подход к обнаружению ухудшения точности устанавливается посредством нахождения корреляции между электрическим сигналом, измеренным подсистемой мониторинга робота, и значениями точности позиционирования (рисунок 10). Разработка регрессионной модели позволила достичь минимальной ошибки при прогнозировании погрешности позиционирования.

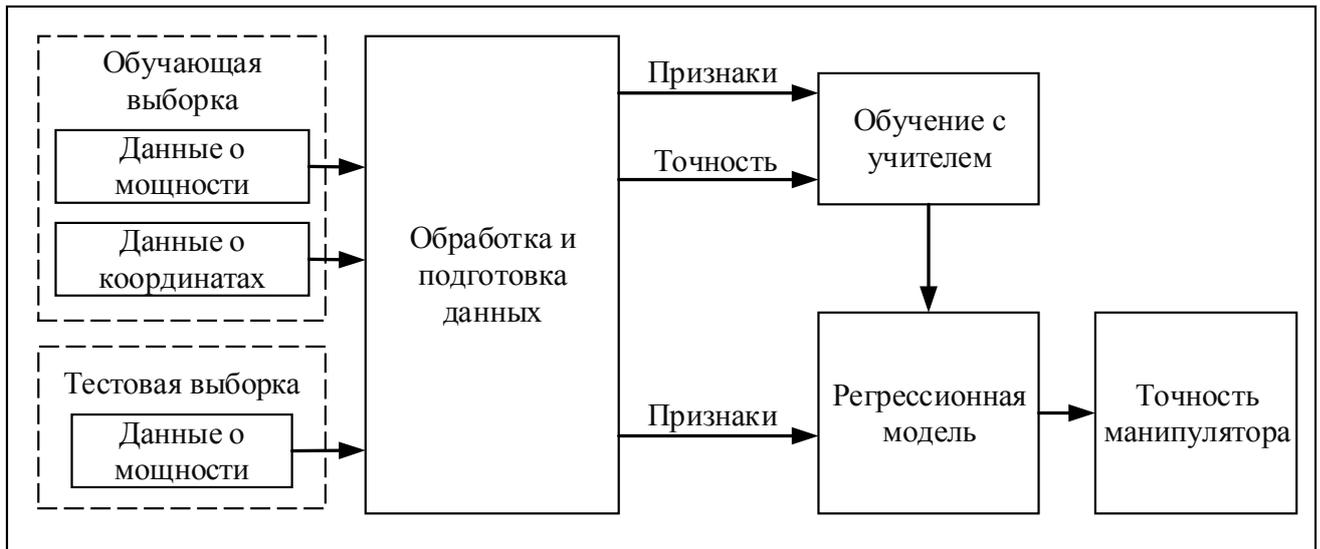


Рисунок 10 – Подход прогнозирования целевого параметра роботизированного манипулятора

В работах [94-96] исследованы подходы и модели управления роботизированными системами с применением методов машинного обучения и ИНС. Обозначенными проблемами являются: недостаток методов построения моделей для самообучения робота в условиях неопределенности и недостаток средств мониторинга состояний робота, приводящих к браку изделия. Как правило, под недостатком «средств мониторинга» подразумевается не отсутствие систем фиксации параметров при рабочем процессе оборудования, а отсутствие имплементации методов анализа и прогноза исторических данных в штатных системах соответствующего вендора. Методами решения обозначенных проблем являются: моделирование деградации оборудования в течении рабочего процесса, корреляционный анализ между телеметрическими данными и оценкой дискретного состояния, машинное обучение (алгоритмы регрессии и классификации), применение метода Уэлча для цифровой обработки временных рядов.

Прогностические методы различаются в зависимости от типа рассматриваемой системы. Классификация систем основывается на том, является ли ее состояние наблюдаемым непосредственно или косвенно, а также моделируется ли ее состояние как дискретный процесс или непрерывный процесс [108]. На рисунке 11 приведено дерево решений данной классификации.

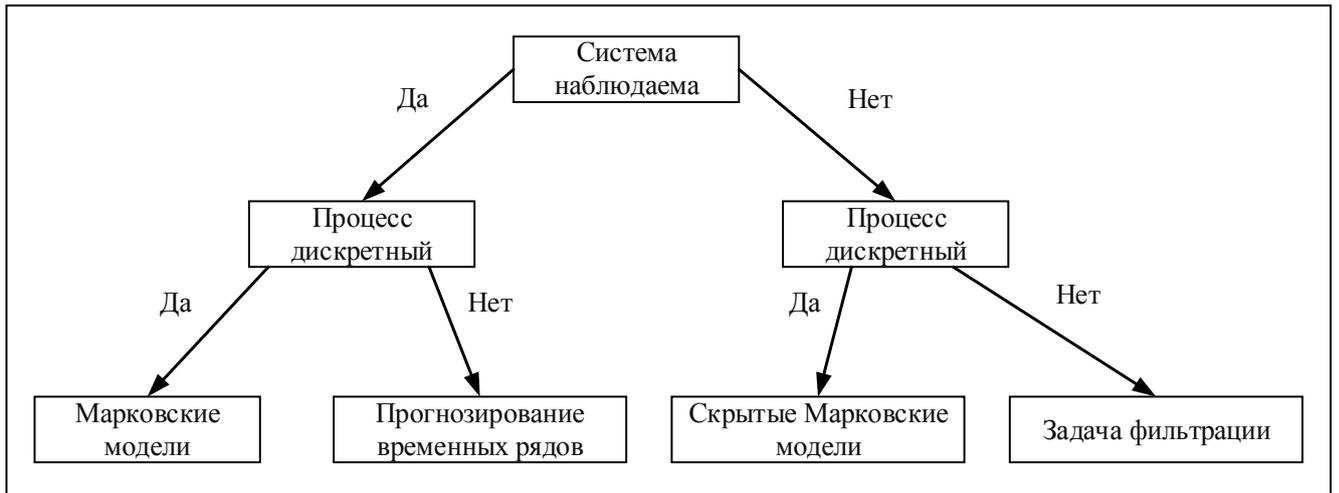


Рисунок 11 – Классификация моделируемой системы

В некоторых случаях отслеживаемые данные могут напрямую описывать базовое состояние системы, таких как износ и размер трещин. В этих случаях оценку RUL можно эффективно сформулировать как задачу прогнозирования временных рядов. Однако во многих других случаях отслеживаемые данные могут лишь косвенно указывать на базовое состояние системы, например, мониторинг вибрации и напряжения для движущихся агрегатов. В данных ситуациях прогностическая задача формулируется как решение двух уравнений:

$$Y_k = h(X_k) + v_k, \quad (6)$$

$$X_k = f(X_{k-1}) + \omega_k, \text{ где} \quad (7)$$

k – временной интервал,

Y_k – измеренные значения признаков,

X_k – внутреннее состояние системы,

$h(X_k)$ – модель измерения,

v_k – шум измерения,

$f(X_{k-1})$ – модель деградации,

ω_k – неопределенность модели.

В некоторых случаях делается предположение, что система развивается в конечном пространстве состояний $\Phi = \{0, 1, \dots, N\}$, где 0 соответствует идеальному исправному состоянию, а N представляет неисправное состояние. Данные дискретные состояния могут быть получены на основе рабочих условий, классифицируемых как «Хорошо», «Только незначительные дефекты», «Требуется техническое обслуживание», или они могут быть получены путем применения методов неконтролируемой кластеризации к обучающим данным. В иных случаях имеет смысл моделировать эволюцию системы как непрерывный процесс.

В работе [97] исследованы подходы к прогнозированию состояний промышленного манипулятора с использованием различных методов машинного обучения. Отмечена проблема в недостатке средств мониторинга критических состояний, приводящих к поломке робота, а также проблема выбора и управления моделью машинного обучения классификации и прогнозирования состояния роботизированной системы. Решением обозначенной проблемы управления моделями является внедрение подхода по автоматическому машинному обучению (AutoML) [98]. Применение AutoML имеет широкое использование, в частности, создание моделей для раннего определения дефектов материалов [99]. В исследовании [100] приводится адаптационный подход для применения автоматического машинного обучения применительно к методу RUL. Данный подход является комплексным решением для автоматической оценки RUL с выбором алгоритма и оптимизацией гиперпараметров модели.

Обозначенная выше проблема по управлению моделями анализа данных исследована в работе [101]. В статье обозначается проблема недостатка средств управления и учета моделями анализа данных, а также приводится перечень актуальных вызовов для исследований процессов управления и оптимизации ресурсов в обозначенном направлении. Частично данная проблема решается через применения подхода DevOps (методология автоматизации технологических процессов сборки, настройки и развёртывания программного обеспечения). Одним из результатов развития направлений DevOps и ML является MLOps – набор практик, нацеленных на надежное и эффективное развёртывание и поддержание

моделей машинного обучения на производстве [102, 103]. Обзорное исследование [104] обобщает выявленные проблемы по управлению моделями анализа данных на каждом из этапов жизненного цикла. В качестве решения проблем на этапе выбора модели, помимо методов выбора алгоритма и оптимизации параметров, предлагается использование подхода трансферного обучения – исследовательская проблема в машинном обучении (ML), которая фокусируется на сохранении знаний, полученных при решении одного кейса, и применении их к другой, но связанной проблеме [105-107].

Одним из подходов автоматизации передачи данных и конструирования моделей анализа данных является Pipeline. Данный подход предполагает организацию информационной инфраструктуры в виде непрерывного конвейера последовательного передачи данных между разными узлами и уровнями. Machine learning pipeline – это последовательное сочетание различных этапов и процессов, предназначенных для разработки, обучения, оценки и развертывания моделей машинного обучения (включая методы MLOps). Архитектура такого конвейера включает в себя все этапы жизненного цикла модели анализа данных: от сбора данных до развертывания и обновления модели [148, 149].

Внедрение подхода конвейеризации данных позволяет сократить ресурсы на стандартизацию, моделирование и обновление программного обеспечения анализа данных (в том числе: оптимизацию и настройку моделей анализа и принятия решений на основе экономических метрик окупаемости затрат предиктивного технического обслуживания) [150].

В исследовании [108] предлагается совместное использование подходов AutoML и трансферного обучения для реализации стратегии обработки данных и управления моделями анализа данных (отбор, дообучение и адаптация с учетом новых источников данных) для значительного сокращения времени разработки, с целью создания конкурентоспособных базовых моделей оценки работоспособности на примере прогнозирования отказов роликов подшипников. Данные, собираемые с акселерометров, стандартизируются и кластеризуются в зависимости от рабочего состояния оборудования. Далее производится выбор из

предварительно обученных и подготовленных моделей для трансферного обучения параллельно с созданием набора базовых моделей из источника автоматического машинного обучения. После отбора лучших базовых моделей производится объединение с трансферной моделью (с использованием ансамблевых методов) для последующей классификации состояния оборудования.

В исследовании [109] рассмотрена проблема оптимизации временных и вычислительных ресурсов при использовании подхода автоматического машинного обучения. Для решения обозначенной проблемы применен подход генетического алгоритма. Генетический алгоритм – это метод поиска, используемый для решения задач оптимизации и моделирования путём случайного подбора, комбинирования и вариации искомых параметров с использованием механизмов, аналогичных естественному отбору в природе. Применительно к задаче отбора алгоритмов машинного обучения, проводится исследование пространства конфигурации, включающей в себя параметры и настройки для управления в рамках предиктивного подхода.

Практическое применение генетических алгоритмов в AutoML [110] включает:

- Эффективный поиск оптимальных конвейеров машинного обучения, сокращающий время и усилия, необходимые специалистам по обработке данных.
- Адаптация к динамичной среде и дрейфу данных, гарантируя, что модели останутся актуальными и точными с течением времени.
- Облегчение сравнительного анализа различных методов AutoML, позволяющего принимать обоснованные решения об использовании конкретного подхода.

Определение трансферного обучения дается с точки зрения областей и постановки задач. Домен D состоит из: функционального пространства X и предельного распределения вероятностей $P(X)$, где $X = \{x_1, \dots, x_n\} \in X$. Принимая в расчет конкретный домен $D = \{X, P(X)\}$, задача состоит из двух компонентов: пространства меток Y и объективная предсказательная функция $f: X \rightarrow Y$. Функция f используется для предсказания соответствующей метки $f(x)$ нового

экземпляра x . Данная задача, обозначенная как $T = \{Y, f(x)\}$, изучается из обучающих данных, состоящих из пар $\{x_i, y_i\}$, где $x_i \in X$ и $y_i \in Y$. Учитывая исходный домен D_S и задачу T_S , целевой домен D_T и задача T_T , где $D_S \neq D_T$ или $T_S \neq T_T$, трансферное обучение направлено на улучшение обучения целевой прогностической функции $f_T(\cdot)$ в D_T , используя знания в D_S и T_S .

Одним из ключевых преимуществ трансферного машинного обучения является возможность использования знаний, полученных при решении одной задачи, для решения другой задачи. Знания, полученные из предыдущих задач, помогают заполнить пробелы в разреженных данных, что способствует более эффективному обучению модели. Трансферное обучение позволяет сократить объем данных, необходимых для обучения модели – вместо обучения модели с нуля на большом наборе данных, можно использовать предобученную модель и дообучить ее на относительно небольшом наборе данных, специфичных для новой задачи. Данный подход экономит вычислительные мощности и время, необходимые для обучения модели с нуля, и позволяет достичь хороших результатов даже при ограниченных ресурсах.

1.4. Выводы по ГЛАВЕ 1

Проведенный анализ актуальных подходов и методов решения задач оценки работоспособности и состояния промышленных роботов позволяет сделать следующие выводы:

1) Сложилась устойчивая тенденция к переходу от профилактической и проактивной стратегии обслуживания к предиктивной, которая позволяет сокращать производственные издержки опираясь на разработанные модели оценки остаточного срока службы оборудования, компонентов и оценку работоспособности в режиме реального времени.

2) Методы интеллектуального анализа данных для стохастического (Марковские цепи, прогнозирование временных рядов, RNN, TCN) и дискретного прогнозирования (задача фильтрации, скрытые Марковские модели, pattern

recognition), а также подхода Process Mining, применение которого позволяет абстрагироваться от низкоуровневых взаимозависимостей в исходных исторических данных до процесса управления данными и моделями анализа данных с целью принятия управленческих решений. Обозначены проблемы при управлении жизненным циклом и оптимизации использования моделей анализа данных.

3) Одним из основных подходов к интеллектуализации процессов оценки состояния промышленных роботов является создание проблемно-ориентированных конвейеров обработки данных, включающих трансферное и автоматическое машинное обучение, для решения выше обозначенных проблем. Представленные результаты анализа позволяют сделать вывод о перспективности исследования и использования методов интеллектуального анализа данных для предиктивного обслуживания оборудования.

4) Несмотря на достигнутый прогресс в исследовании проблем оценки работоспособности и состояния промышленных роботов, частично решенными остаются такие задачи как: разработка методов выбора оптимальных прогнозных моделей интеллектуального анализа данных, дообучение и переобучение прогнозных моделей, а также управление процесса поддержки принятия решений по стратегическому планированию технического обслуживания на основе данных.

Таким образом, изложенные выше выводы позволяют сформулировать основные задачи диссертационного исследования:

- Разработка архитектуры программного обеспечения с использованием проблемно-ориентированного конвейера обработки данных, допускающую возможность расширения, тиражирования и адаптивности под специфику конкретного производственного процесса для обеспечения поддержки принятия решений на уровне стратегического управления и планирования технического обслуживания разнородного производственного оборудования.
- Разработка методов интеграции разнородных данных, собранных с парка промышленных роботов, извлечения знаний из существующего объема

информации, с учетом сложности и эффективности обработки имеющейся информации.

- Разработка подходов выбора оптимальных прогнозных моделей для анализа данных разнородного промышленного оборудования, с учетом оптимизации гиперпараметров моделей машинного обучения, оптимизации структур нейронных сетей, а также минимизацией вычислительных и временных ресурсов на обучение моделей.

- Применение методов оценки уровня точности прогнозирования и экономической целесообразности для принятия решений по техническому обслуживанию на основе использования метрик прогнозных моделей.

2. Проблемно-ориентированный конвейер обработки данных для прогнозирования работоспособности промышленных роботов с использованием методов интеллектуального анализа

2.1. Формальное представление процесса принятия решений ТОиР в рамках предиктивного подхода

Управление жизненным циклом моделей анализа данных (в частности – машинного обучения и искусственных нейронных сетей) является одной составляющих оптимизации временных и вычислительных ресурсов в процессе использования моделей для систем поддержки принятия решений [96].

Разработка архитектуры для реализации анализа исторических данных о работе основных средств производства производится в соответствии с ограничениями процесса планирования технического обслуживания [134]. Данный процесс включает в себя следующие операции:

1) Сбор и хранение данных из разных источников: датчиков, контроллеров, систем мониторинга.

2) Анализ данных:

2.1) Подготовка данных:

2.1.1) Профилирование данных.

2.1.2) Сбор данных.

2.1.3) Очистка данных.

2.1.4) Моделирование данных.

2.1.5) Конструирование признаков.

2.1.6) Документирование данных.

2.1.7) Доставка данных.

2.2) Моделирование:

2.2.1) Выбор алгоритма.

2.2.2) Планирование тестирования.

2.2.3) Обучение модели.

2.2.4) Оценка результатов.

3) Принятие решений:

3.1) Анализ результатов тестирования.

3.2) Классификация состояния и оценка работоспособности оборудования.

3.3) Корректировка плана технического обслуживания производственного оборудования.

3.4) Назначение процедур по техническому обслуживанию производственного оборудования в соответствии со скорректированным планом.

Декомпозиция данного процесса представлена с использованием нотации IDEF0 на рисунках 12–13. Более подробная декомпозиция анализируемого процесса, включающая в себя этапы подготовки данных, создания и обучения моделей интеллектуального анализа, представлена на рисунках 56–59 (Приложение 3).

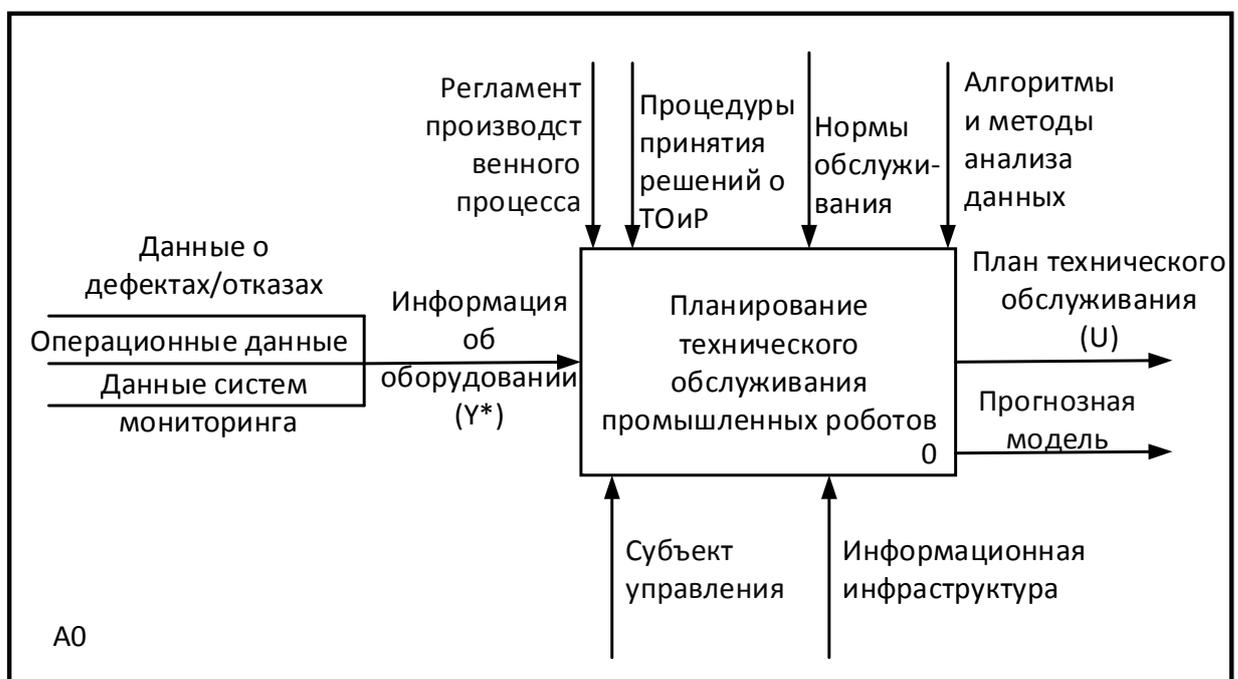


Рисунок 12 – Процесс анализа данных для планирования ТОиР в нотации IDEF0 (блок A0)

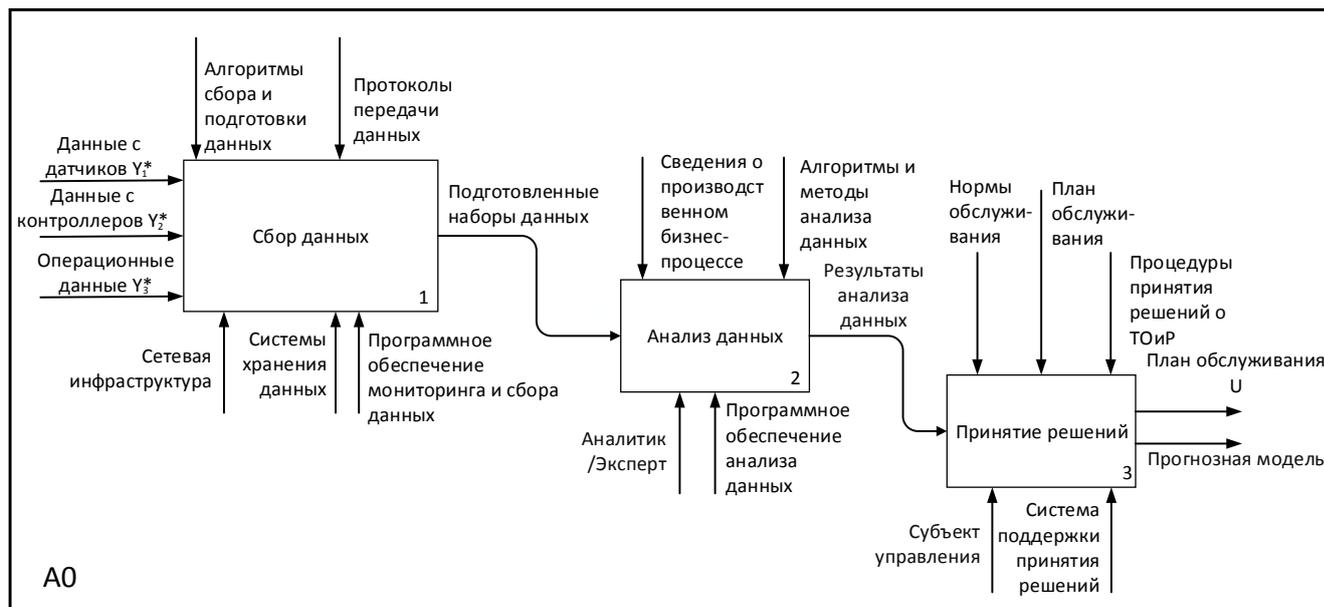


Рисунок 13 – Процесс анализа данных для планирования ТОиР в нотации IDEF0 (декомпозиция блока А0)

2.2. Методика проектирования проблемно-ориентированного конвейера обработки данных, обеспечивающего реализацию процесса принятия решений ТОиР на основе данных

Для реализации систем поддержки принятия решений планирования технического обслуживания и оценки состояния необходимо наличие следующих составляющих [134, 155]:

- Подсистема сбора данных: обеспечение доступа к получению информации с цифровых датчиков, преобразователей данных, контроллеров и систем мониторинга.
- Подсистема управления данными: алгоритмическое обеспечение для доступа к массивам собранных данных.
- Подсистема мониторинга состояния: обеспечение мониторинга текущего состояния оборудования – сравнение данных мониторинга с ожидаемыми значениями или пороговыми значениями, включая функцию оповещения при возникновении ошибок/нештатных ситуаций.

- Подсистема оценки работоспособности промышленного оборудования: обеспечение анализа тренда на уменьшение срока полезной эксплуатации.
- Подсистема прогностической оценки: обеспечение оценки будущего состояния работоспособности оборудования.
- Подсистема поддержки принятия решений: обеспечение рекомендаций для планирования технического обслуживания и настройке конфигурации производственного оборудования с учетом исторических данных эксплуатации, текущих и будущих сценариев использования и ограничениями по обеспечению ресурсами для работы и обслуживания.

В работе [136] рассмотрен вопрос по проектированию интеллектуальной системы мониторинга и диагностики для разрешения проблемных ситуаций, посредством своевременного предупреждения оператора о сбоях оборудования, что позволяет сократить материальный и экологический ущерб. Ключевыми особенностями предложенной архитектуры являются: применение паттернов проектирования «Наблюдатель» и «Правило», а также подхода оценки матричного и мультиматричного профиля. Матричный профиль основан на одном классе наблюдений, а мультиматричный профиль основан на нескольких классах наблюдений [137]. Использование данного подхода позволяет выявлять аномальные значения во наборах данных, оценивать сходство временных рядов для поиска типичного паттерна или редкого «выброса», что позволяет провести оценку и прогноз состояния оборудования.

Отдельными элементами по работе с данными являются программные модули, отвечающие за разметку, обработку и поиск аномальных значений. Аномальные значение могут появляться на различных этапах обработки данных – от сбора (в силу объективных технических причин) до анализа (с целью выявления интересующих субъекта управления знаний). Исходя из этого можно выявить следующие ограничения в процессе выявления аномалий:

- Неизвестность новых данных – методы поиска аномалий должны обобщать предыдущие результаты анализа и выявлять схожесть в анализируемых данных.
- Гетерогенность данных – в условиях анализа данных производственной линии разнородного оборудования, непохожесть данных может вносить существенные помехи в процесс анализа.
- Частота появления аномалий – редкость и нерегулярность (т.е. отсутствие трендов) может затруднять обучение моделей с учителем даже в условиях корректной разметки данных.
- Обширность и непредсказуемость аномальных явлений.

В контексте методов анализа больших данных можно выделить следующие алгоритмы для выявления аномальных значений: одноклассовая классификация (One-Class SVM [160], isolation forest [161]), кластерный анализ (DBSCAN, ABOD [162], LOF [163]), глубокое обучение (Autoencoder [164], GAN [165], Anomaly Measure-dependent Feature Learning).

В работе [142] приводится двух-кольцевая архитектура для управления жизненным циклом моделей анализа данных в рамках прогнозного технического обслуживания ветряных турбин. Ключевой особенностью предлагаемой архитектуры является обучение и настройка моделей в два этапа: после сбора и подготовки данных и после извлечения значимых атрибутов для выявления факторов, оказывающих наибольшее влияние на предсказательную точность модели (рисунок 14). Отброс неинформативных признаков позволяет оптимизировать временные ресурсы на обучение модели и повысить точность.

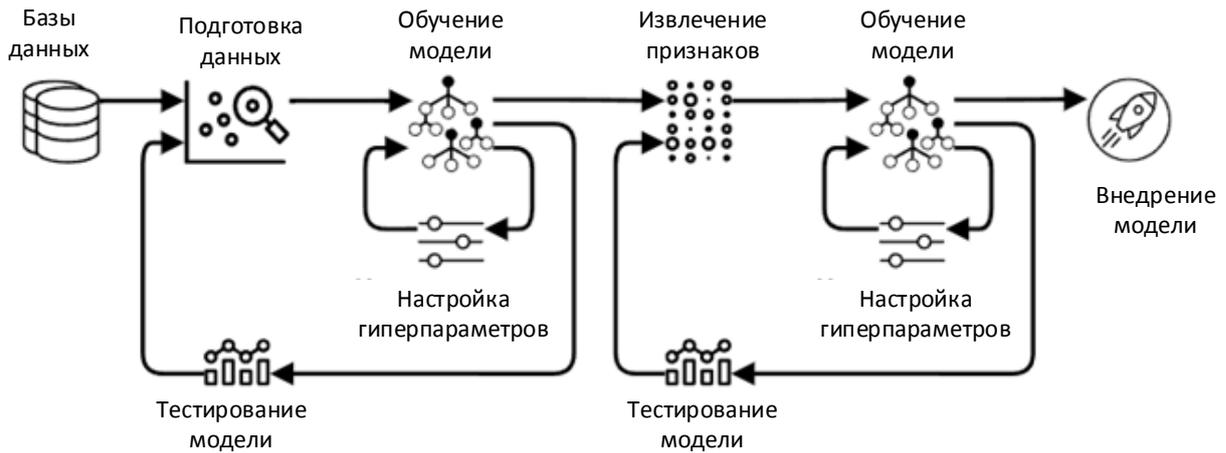


Рисунок 14 – Архитектура двухэтапного обучения модели анализа данных

В работе [142] приводится пример разработки мультисенсорной архитектуры анализа данных в рамках подхода федеративного машинного обучения. Федеративное обучение – это метод распределенного машинного обучения, который позволяет обучать модели на нескольких устройствах или серверах без обмена образцами данных. При федеративном обучении каждое устройство или сервер собирает и обрабатывает данные локально и отправляет только обновления модели на центральный сервер, который объединяет их для улучшения метрик итоговой модели [143].

В техническом обзоре [151] исследуются составляющие подхода MLOps, включающие в себя: компоненты непрерывной доставки и развертывания (CI/CD), системы контроля версий, компоненты оркестровки рабочего процесса (на примере Pipeline), системы генерации моделей анализа данных, сервера для обучения (которые предоставляют основные вычислительные мощности), хранилища и реестры моделей интеллектуального анализа [153]. Проектирование информационной инфраструктуры для реализации такого подхода требует высокой квалификации специалистов по следующим направлениям: архитектор программных решений, Data Science, Data Engineering, DevOps и MLOps Engineering. Оркестровка и конвейеризация управления данными включают координацию различных этапов для процесса обработки данных и обучения моделей. Данный подход включает в себя следующие основные шаги:

1. Сбор данных из различных источников, таких как базы данных, API, файловые системы и пр.
2. Подготовка данных – удаление выбросов, заполнение пропущенных значений, кодирование категориальных признаков и масштабирование данных.
3. Обучение моделей – выбор подходящих моделей, настройку гиперпараметров и оценку их производительности.
4. Оценка моделей – использование метрик качества для определения их эффективности.
5. Развёртывание моделей – модели развёртываются в производство с помощью сервисов облачных платформ или инструментов контейнеризации.
6. Мониторинг и обновление – периодическое обновление для улучшения их производительности на основе новых данных или изменяющихся требований.

Автоматизация этого процесса осуществляется с помощью инструментов для управления рабочими процессами, таких как Apache Airflow, Kubeflow Pipelines, MLflow, Luigi и др. Для автоматизации процессов управления данными, моделями и результатами анализа предложена методика организации программного обеспечения на основе подхода проблемно-ориентированного конвейера. В отличие от информационных систем, построенных на монолитном подходе (высокая зацепленность программных модулей и слабая масштабируемость), событийно-ориентированном (повышенные требования к контролю за процессами управления событиями) и на базе задач (необходимость в отслеживании зависимостей для обеспечения корректной последовательности выполнения). В общем виде система автоматизации управления и мониторинга моделями машинного обучения представлена на рисунках 15–16.

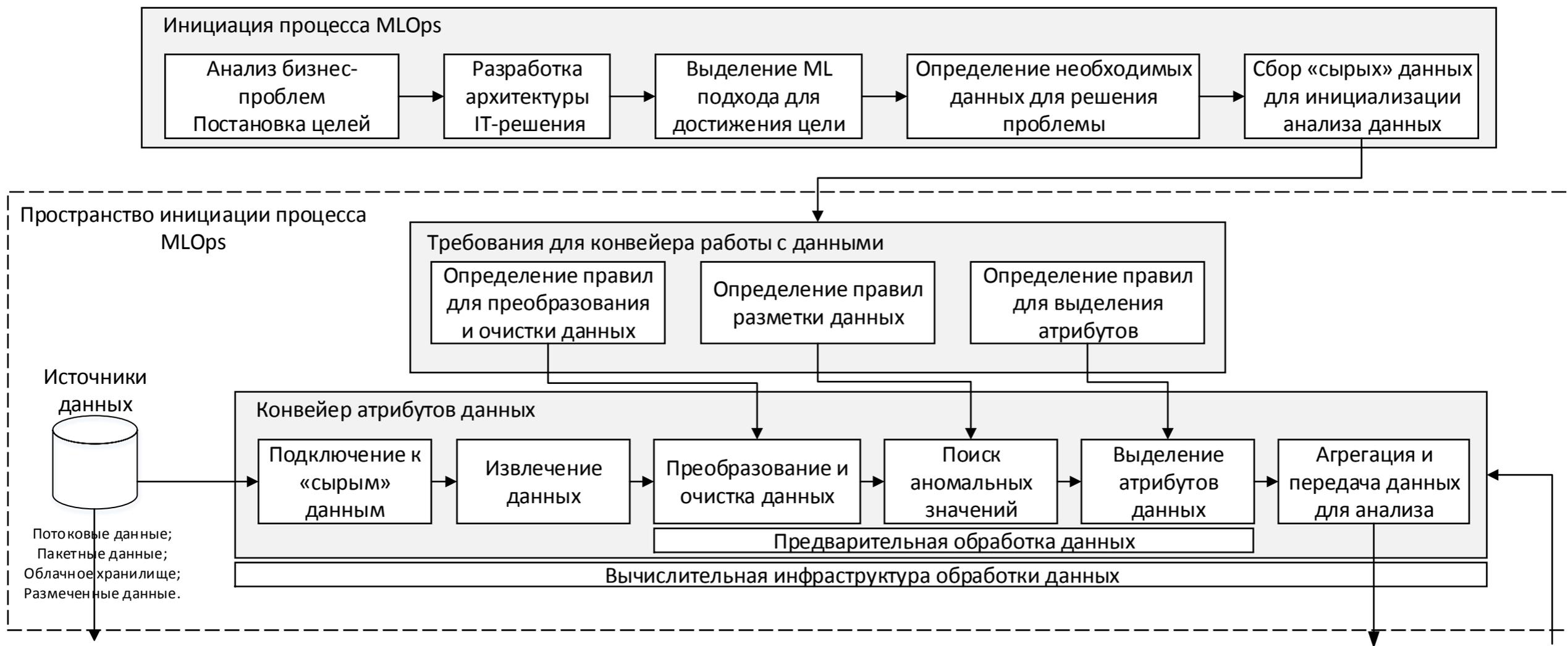


Рисунок 15 – Процесс автоматизации подготовки данных и построения моделей машинного обучения

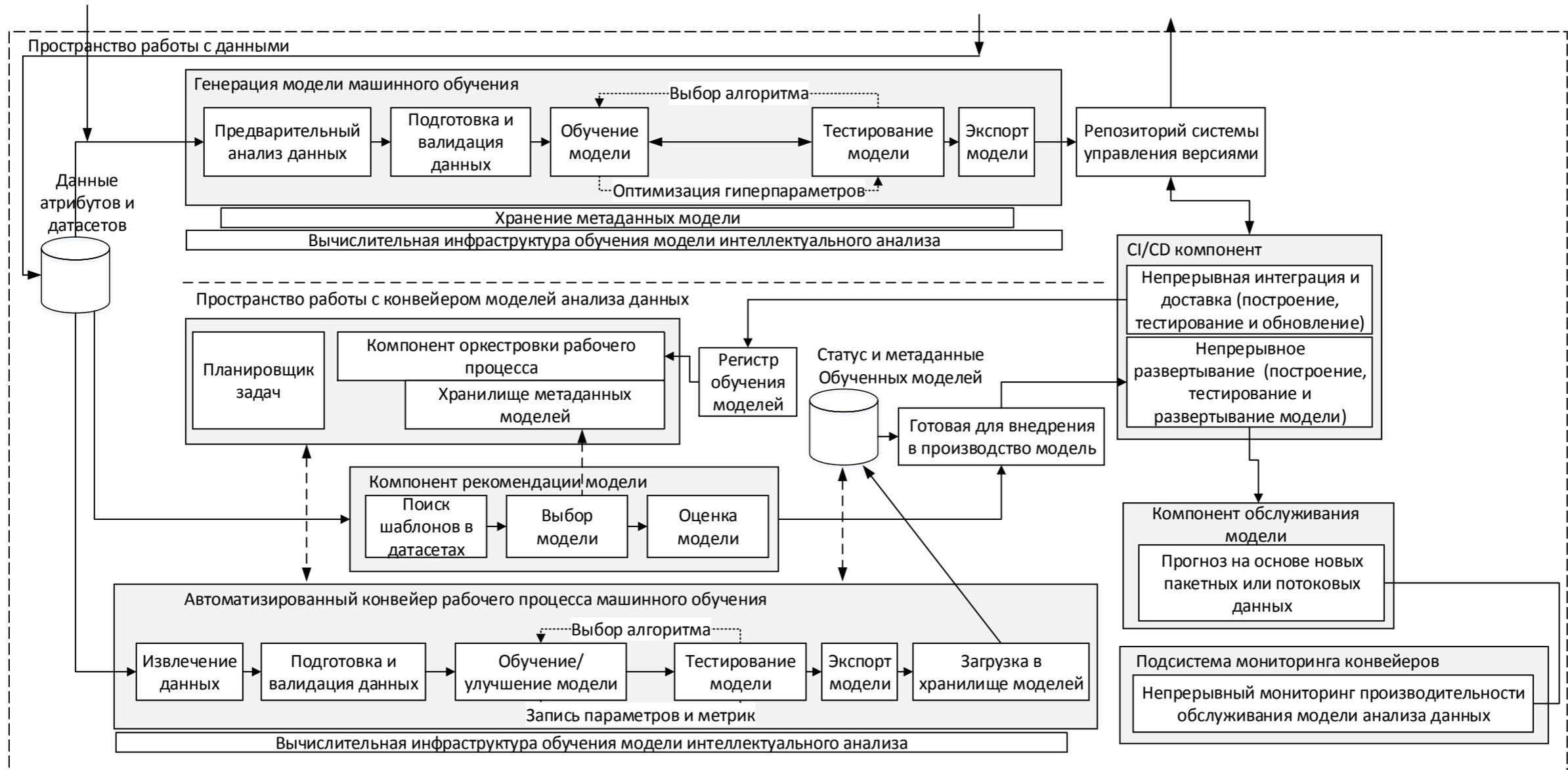


Рисунок 16 – Процесс автоматизации подготовки данных и построения моделей машинного обучения

В исследовании [152] произведен обзор технологий построения и оркестровки процесса управления моделями, а также приведен анализ 3000 конвейеров генерации и обучения моделей машинного обучения с целью поиска «узких мест» и возможности для оптимизации временных и вычислительных затрат. Авторы предлагают подход по оптимизации конвейерной архитектуры на основе методов управления данными, за счет сокращения избыточных вычислений. Предлагаемый подход по выявлению и сокращению избыточных вычислений, которые не связаны с развертыванием модели, позволяет снизить вычислительные затраты на 50%.

В работе [154] предложен прикладной подход по автоматизации проектирования комплексных конвейеров машинного обучения (которые эквивалентны рабочим процессам), состоящим из моделей и операций с данными. Подход сочетает в себе ключевые идеи как автоматизированного машинного обучения, так и систем управления рабочими процессами. Предлагаемый способ позволяет проектировать конвейеры с настраиваемой графической структурой, анализировать полученные результаты и воспроизводить их. Эволюционные методы используются для гибкой идентификации структуры конвейера, что так же рассмотрено в исследовании [156] на примере различных фреймворков конструирования автоматического машинного обучения.

2.3. Подходы и методы управления моделями анализа данных

Современная техника оснащена достаточным количеством датчиков, которые позволяют отслеживать его работу. Данные, получаемые от датчиков, представляют собой многомерные временные ряды:

$$X = \{x_{t_k}\}_{k=1}^L, \text{ где} \quad (8)$$

L – длина временного ряда,

$x_{t_k} = \{r_i(t_k) | i = 1, \dots, n\}$, $x_{t_k} \in \mathbb{R}^n$ - n -мерный вектор показаний, передаваемых n датчиками в момент времени t_k . Часть показаний $x_{t_k}^{(s)} \in \mathbb{R}^{(s)}$ может поступать с датчиков, собирающих данные с узлов и агрегатов оборудования, часть $x_{t_k}^{(c)} \in \mathbb{R}^{(c)}$ с датчиков или источников, предоставляющих данные об условиях эксплуатации, $n = s + c$. Учет внешних факторов при прогнозировании состояния оборудования позволяет построить более полную модель, описывающую изменение его состояния и работоспособности. Задача предиктивного прогнозирования для поддержки принятия решений при ТОиР заключается в обработке набора данных X в каждый момент времени t_k сопоставляемых с целевыми показателями работоспособности и оценки состояния оборудования Y для последующего принятия решений по формированию и реализации задач технического обслуживания. Разработка комплексных решений для анализа технологических данных связано со следующими сложностями:

- Нестационарность и динамичность технологических процессов, обуславливающие сложность математического описания.
- Неоднородная структура данных (вещественные значения, события, сообщения).
- Обширный диапазон набора собираемых данных, зависящий от оборудования, набора датчиков, фирмы производителя системы мониторинга.
- Уникальность каждого объекта управления (в отличие от биомедицинских и речевых данных, которые имеют обобщённые характерные признаки), которая требует адаптивного подхода для каждого случая.
- Наличие аномальных ситуаций, которые могут являться как следствием изменения интересующего тренда технологического сигнала, так и следствием ремонтных или иных регламентных процедур.

Наличие решений для разрешения такого рода сложностей позволяет извлекать полезную информацию о технических характеристиках агрегатов на производстве, локализовать с максимальной точностью область и причину

неисправности оборудования и прогнозировать состояние объекта с учетом его ретроспективных показателей и характеристик других смежных агрегатов [112].

2.3.1. Подходы по конфигурации и выбору модели интеллектуального анализа данных

Обучение, техническое обслуживание, развертывание, мониторинг, организация и документирование моделей машинного обучения – являются важнейшими задачами практически во всех случаях использования машинного обучения в производственной среде. Эффективность применения моделей интеллектуального анализа данных оценивается по метрикам. Для разного класса методов существуют соответствующие наборы метрик, каждая из которых имеет критерии оценки эффективности применяемого метода для решения поставленной задачи. В общем виде оценку моделей интеллектуального анализа можно описать следующими этапами:

- 1) Подбор параметров и гиперпараметров модели.
- 2) Обучение модели с заданными параметрами.
- 3) Тестирование модели с заданными параметрами.

Гиперпараметры модели – это параметры, которые устанавливаются перед этапом обучения модели [122]. К гиперпараметрам относятся следующие свойства (по категориям моделей):

- 1) Искусственные нейронные сети:
 - Крутизна активационной функции.
 - Число нейронов и скрытых слоев.
 - Коэффициент скорости обучения.
 - Число итераций обучения.
 - Уровень ошибки обучения.
- 2) Деревья решений:
 - Минимальное число обучающих примеров.

- Максимальная глубина дерева.

3) Кластеризация:

- Число кластеров.
- Начальный радиус обучения.

4) Ассоциативные правила:

- Значения поддержки и достоверности правил.

5) Модели регрессии:

- Дискриминационный порог.

Определение фактической модели управления является нетривиальной задачей. В узком смысле параметры рассматриваемой модели, полученные после обучения (например, набор весов модели логистической регрессии), могут рассматриваться как математическая модель для управления. Однако существует необходимость преобразования входных данных в функции, которые являются входными для модели. Соответствующие преобразования функций являются важной частью модели, которыми также необходимо управлять. Существуют программные библиотеки для управления конвейерами ML (Scikit-Learn, Apache Spark), которые объединяют преобразования функций и реальную модель ML в рамках одного абстрактного решения. В работе [130] представлен инструмент для управления моделями машинного обучения. Управление производится с помощью библиотек: Spark (создание конвейера обработки данных и обучения модели), Scikit-Learn (генератор методов машинного обучения и тестирования), R (создание и использование конфигурации модели). Системы отслеживания и хранения моделей, которые позволяют связать преобразования признаков с параметрами модели представлены в исследовании [131]. Из-за предположения независимости и равномерно распределенных данных, присущего многим алгоритмам ML, модели содержат неявные предположения о распределении данных, к которым они применяются. Нарушение этих предположений (например, ковариантный сдвиг данных) может привести к снижению качества прогноза. Захват и управление этими неявными предположениями требуют разработки системных подходов [132, 133]. Существует еще один аспект в управлении моделями – представление

отдельных моделей в виде «черных ящиков», то есть абстрагирование от алгоритма работы модели до входных и выходных данных, которые обеспечивают результат анализа. В рамках конвейерного подхода для управления «черными ящиками» используется метод «рабочего процесса» (workflow). Для реализации данного метода существует программное обеспечение, позволяющее проектировать рабочий процесс без программирования, то есть с использованием визуального представления процесса в виде схемы с отдельными элементами и связями между ними, которые определяют потоки данных и интерфейсы взаимодействия.

После этапа построения модели, выполняется первичная валидация для оценки работоспособности и целесообразности внедрения разработанной модели. На этапе внедрения проводится два вида проверок:

- 1) Верификация – подтверждение качества модели на актуальном потоке данных и дополнительная проверка репрезентативности данных, использованных при разработке модели.

- 2) IT-валидация – аудит набора скриптов с реализацией модели посредством проверки кода на обработку пропусков, дубликатов и других артефактов данных для снижения риска некорректного поведения модели.

Валидация моделей машинного обучения представляет из себя комплексный процесс, осуществляемый на протяжении всего жизненного цикла модели. Декомпозиция цикла валидации представлена на рисунке 17.

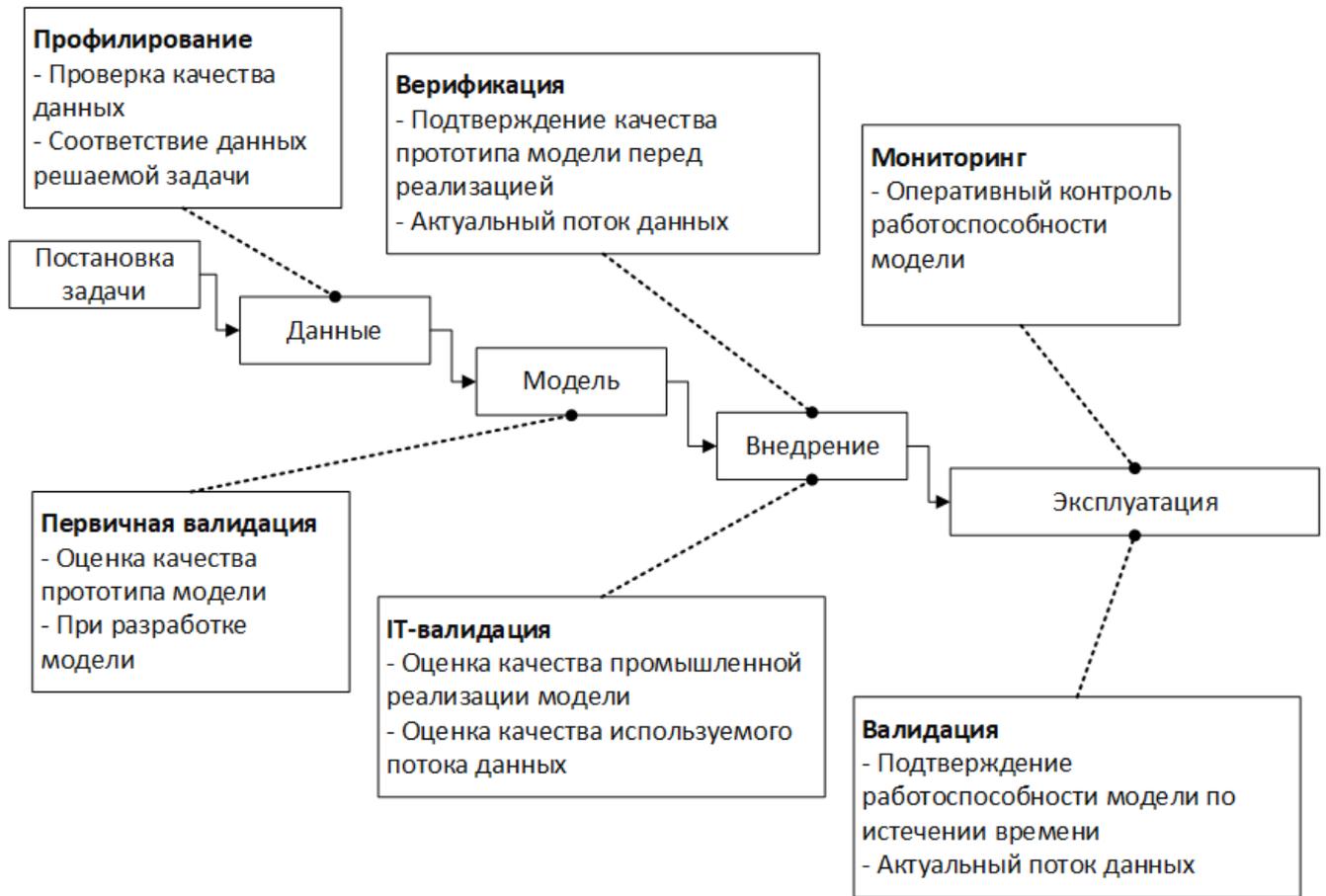


Рисунок 17 – Валидация модели анализа данных на этапах жизненного цикла

На этапе конструирования и валидации модели для прогностического технического обслуживания промышленного оборудования учитывается следующая специфика: наличие телеметрических данных о работе оборудования и исторических данных неисправностей, работ технического обслуживания и регламента обслуживания (рисунок 18).



Рисунок 18 – Процесс проведения прогностического ТО на основе анализа данных

Переобучение моделей машинного обучения представляет собой один из важных этапов управления, обеспечивающий наиболее эффективный выбор параметров настройки моделей для решения поставленной задачи. Обучение проводится в автономном режиме, а модели загружаются во время прогнозирования в обслуживающей системе. Если происходят новые события, на которых текущие модели не обучались, необходимо инициировать переобучение или повторное моделирование. Стандартные методы перекрестной проверки помогают оценить ошибку обобщения и настроить пороги точности для моделей.

2.3.2. Методы отбора значимых признаков

В статье [113] рассмотрен алгоритм исключения малозначимых и сильно связанных признаков. Предложенный алгоритм использует в основе методы генетического алгоритма, методы кластеризации, что обеспечивает прирост точности по сравнению с другими эвристическими методами.

В исследовании [114] произведено описание метода параллельной обработки признаков с использованием технологии MapReduce для выбора оптимального набора входных признаков (с целью уменьшения временной сложности). Произведен сравнительный анализ используемых методов обработки больших наборов данных (PSO, APSO и ASAMO) и оценка точности EIDMLP с алгоритмами классификации данных (Наивный Байес, k-средних и т.д.), показавший большую скорость и точность по сравнению с аналогами.

В работе [116] предложен метод неконтролируемого отбора признаков. Рассматриваемая проблема: высокая сложность отбора признаков для сокращения времени вычислений. Алгоритм посылает более высокую точность по сравнению с другими методами контролируемого, неконтролируемого отбора признаков.

В исследовании [117] описывается метод выбора признаков из четких и неопределенных наборов данных. Предложенный метод направлен на уменьшение количества используемых признаков модели; основан на использовании Fuzzy Random Forest [118], Filter и Wrapper методах.

В работе [119] разработан метод выбора признаков для систем принятия решения в основе которых лежат большие объемы данных. Предложенный метод комбинирует использования грубых методов оценки неопределенности, выявление связанности признаков и эвристических алгоритмов низкой вычислительной сложности для повышения точности выбора признаков [120].

В работе [121] описывается подход отбора наиболее актуальных и влиятельных признаков. В качестве решения предложена комбинация методов для отсеивания малозначимых признаков (FMCCSC и KNN).

2.3.3. Методы оценки эффективности моделей интеллектуального анализа данных

Метрики модели – вычисляемые значения, позволяющие определить эффективность применения модели для решения поставленной задачи [123]. Использование метрик также требует введения критерия для «успешности» модели, позволяющего сравнивать разные модели и делать вывод об оптимальности выбранного алгоритма и/или решения. Для выполнения тестирования моделей в рамках диссертационного исследования выбраны следующие метрики:

1) Классификация:

- Точность (accuracy).
- Правильность (precision).
- Полнота (recall).
- F1-оценка (f1-score).

2) Регрессия:

- Средняя абсолютная ошибка (Mean Absolute Error).
- Средняя квадратичная ошибка (Mean Squared Error).
- Время обучения/тестирования модели машинного обучения.

Задача выбора гиперпараметров обучения заключается в подборе таких параметров, при которых заданная модель выбранного алгоритма обучения будет наиболее эффективна. Гиперпараметры могут выбираться из ограниченного множества или с помощью перебора из неограниченного множества гиперпараметров – это зависит от специфики решаемой задачи. Для выполнения задачи определения оптимального набора гиперпараметров обучения модели машинного обучения выбраны следующие методы:

- 1) Метод «ручного» перебора.
- 2) Метод случайного поиска.

На рисунке 19 представлена диаграмма процесса настройки и обучения модели на основе метрик тестирования.

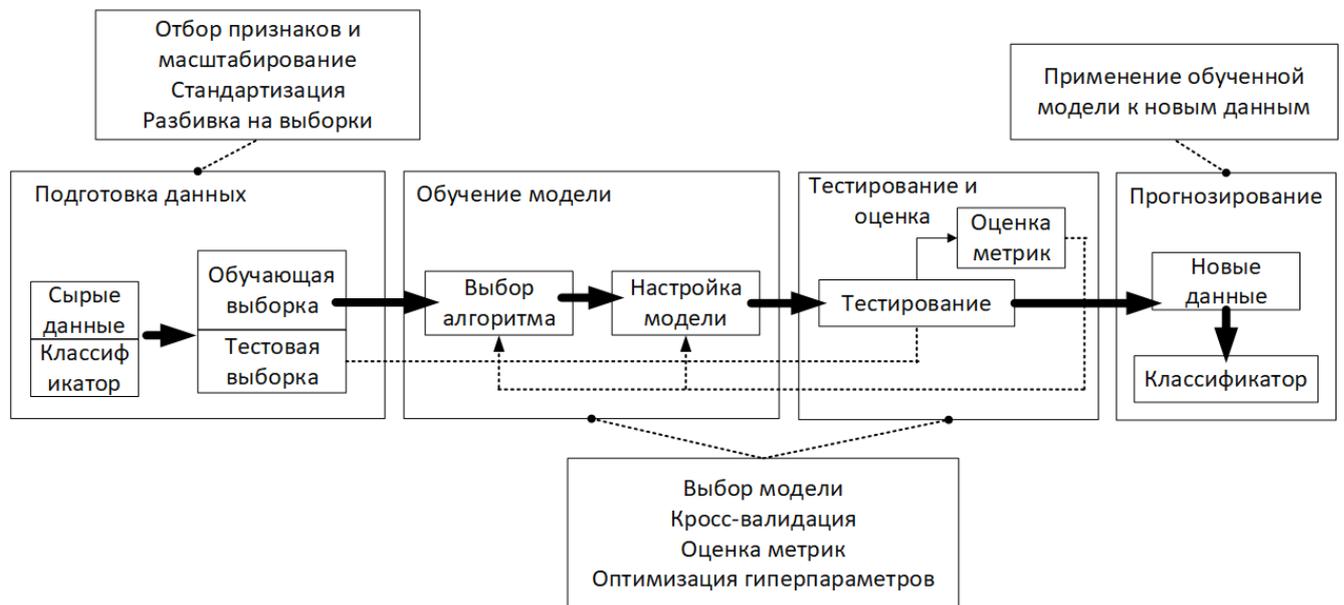


Рисунок 19 – Настройка и обучение модели на основе метрик тестирования

Пример эффективного применения метрик представлен в работе [75] по прогнозированию точности позиционирования промышленного манипулятора – использованы следующие метрики для оценки эффективности модели:

- Среднеквадратичное отклонение.
- Коэффициент асимметрии.
- Коэффициент эксцесса.
- Автокорреляция.
- Энергия непрерывного сигнала.

Использование результирующих значений метрик тестирования модели позволило обеспечить обратную связь для дообучения модели при приемлемом уровне погрешности прогнозирования параметров.

2.3.4. Алгоритмы для улучшения показателей моделей интеллектуального анализа данных

Основным подходом улучшения метрик при тестировании модели является применение ансамблевых методов. Данный подход определяет перечень методов, целью которых является увеличение эффективности решаемой задачи за счет объединения моделей, обученных для решения задачи. Данные методы применяются в случаях, когда обученная модель с оптимальным набором параметров не показывает достаточной точности и полноты для решения поставленной задачи. Гипотеза ансамблевых методов состоит в увеличении точности моделей при правильном сочетании «слабых» моделей. Под правильным сочетанием понимается та ситуация, при которой подбираются модели, нивелирующие недостатки друг друга и не оказывающие негативное влияние на сопредельные алгоритмы. В машинном обучении при выборе модели важно учитывать дилемму смещения-дисперсии – это свойство набора моделей предсказания, когда модели с меньшим отклонением от имеющихся данных имеют более высокую дисперсию на новых данных (то есть подвержены переобучению), и наоборот (рисунок 20).

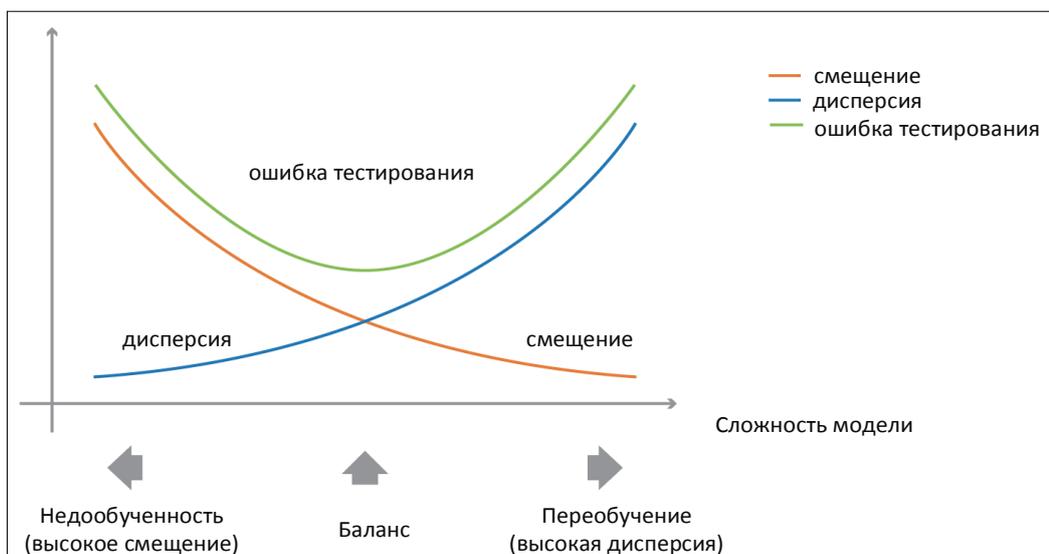


Рисунок 20 – Компромисс смещения-дисперсии

В ансамблевой теории обучения моделей ключевым понятием является «базовая модель» – составляющая блока моделей. Существуют однородные и разнородные ансамблевые модели. Однородная модель агрегирует несколько разных моделей, обученных на основе одного и того же базового алгоритма. Разнородная модель агрегирует несколько разных моделей, обученных на различных типах базовых алгоритмов.

Существуют три основных алгоритма агрегирования базовых моделей [144]:

- 1) Бэггинг (Bagging);
- 2) Бустинг (Boosting);
- 3) Стекинг (Stacking).

Для реализации задач, поставленных в рамках диссертационного исследования, были выбраны следующие подходы:

- Для отбора значимых признаков (методы фильтрации (корреляция Пирсона, дисперсия) и методы обертки (метод последовательного выбора признаков). Следующие подходы целесообразны к использованию при небольшом количестве атрибутов;
- Конвейерная подход по управлению моделями, позволяющий гибко настраивать гиперпараметры и выбор метрик тестирования для реализации автоматического машинного обучения;
- Методы улучшения показателей модели анализа данных (стекинг и бустинг), позволяющих комбинировать архитектуры разных моделей для нивелирования их недостатков и повышения точности прогнозирования.

2.4. Методы оценки сходства данных

Оптимизация использования моделей машинного обучения может достигаться за счет использования моделей на данных, имеющих определенную степень сходства с обучающей выборкой [124]. В общем виде сравниваемые

выборки данных преобразуются в вектора, с последующей оценкой коэффициента сходства или расстояния.

Метод, используемый для измерения расстояний, зависит от конкретной задачи и набора данных. Разные проблемы и задачи требуют могут потребовать более комплексного подхода для расчета расстояний между точками или наблюдениями, например, косинусного расстояния. различные методы вычисления расстояний между каждой парой точек данных:

- Евклидово расстояние.
- Расстояние городских кварталов.
- Расстояние Канберра.
- Косинусное расстояние.
- Корреляционное расстояние Пирсона.
- Корреляция Спирмена.
- Расстояние Жаккара.
- Матричный профиль.

Для эффективного подбора модели машинного обучения необходимо оценить входные данные на подобие с теми данными, которые уже были использованы для обученных модели. Рассмотрим следующие два метода:

- Евклидово расстояние.
- Алгоритм динамической трансформации временной шкалы.

Евклидово расстояние – это фундаментальная метрика расстояния, относящаяся к системам в Евклидовом пространстве. Это кратчайшая прямая между двумя точками в евклидовом пространстве. Евклидово расстояние широко используется в разработке алгоритмов машинного обучения, которые включают в себя метрики расстояния, такие как k -ближайшие соседи. Как правило, Евклидово расстояние будет представлять, насколько похожи две точки данных, предполагая, что некоторая кластеризация на основе других данных уже была выполнена. Евклидово расстояние между точками x и y в n -мерном пространстве вычисляется по формуле, где i – количество координат точки.

Алгоритм динамического трансформирования времени (DTW) вычисляет оптимальную последовательность трансформаций (деформаций) времени между двумя временными рядами [125–127]. В отличие от Евклидова расстояния, алгоритм DTW рассчитывает расстояние между рядами без учета локальных и глобальных сдвигов на временной шкале. Данный метод используется в связке с моделями машинного обучения для сравнения прогнозируемых данных с историческими данными, используемых для обучения, что позволяет снизить ошибку при использовании модели [128, 129].

2.5. Методы оценки экономической эффективности и принятия решений

Принятие решений представляет собой комплексный и нетривиальный процесс, включающий в себя множество шагов и акторов (рисунок 21). В статье [145] рассматриваются основные этапы технологии разработки управленческих решений:

1. Выявление и описание проблемной ситуации – проблема определяется как несоответствие между желаемым и фактическим состоянием объектов управления.
2. Формирование целей системы управления – для определения целевого состояния по устранению проблемной ситуации необходимо сформулировать множество целей системы.
3. Формирование критериев выбора решений – для описания цели должно быть введено столько критериев, чтобы они достаточно полно характеризовали цель при минимальном их числе.
4. Разработка решений – на данном этапе вырабатываются альтернативные варианты решений, осуществляется поиск способов достижения поставленных целей в рамках сформированных критериев.

5. Согласование и выбор решения – на данном этапе необходимо осуществить выбор решения по определенной схеме или алгоритму, наилучшему с точки зрения некоторого критерия, некоторого принципа оптимальности.

6. Реализация и оценка решения – план реализации выбранного решения должен дать ответы на вопросы – кто и что должен делать, какими средствами и в какие сроки. Более подробно методы принятия решений описаны в работе [146].

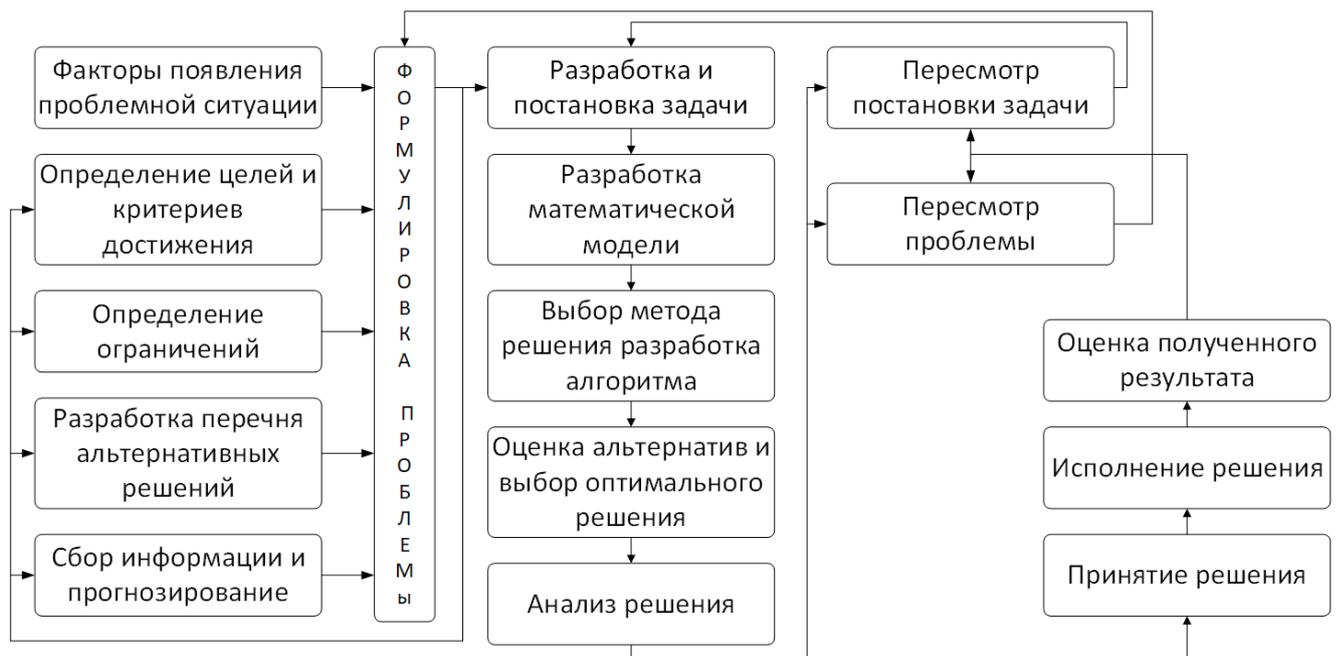


Рисунок 21 – Модель принятия решений на основе данных в общем виде

Метод расчета экономической эффективности для технического обслуживания может варьироваться в зависимости от конкретной ситуации и используемых данных. Однако, в общем виде, можно использовать следующую формулу:

$$k_e = \frac{\Delta P}{C}, \text{ где} \quad (9)$$

k_e – коэффициент абсолютной экономической эффективности,

ΔP – прирост прибыли,

C – объем затрат.

Выгоды включают: увеличение производительности, снижение рисков поломки или простоев оборудования, улучшение качества процессов, сокращение времени простоев и другие факторы, которые приводят к улучшению результатов работы. Затраты включают расходы на обслуживание, запасные части, оборудование, персонал и другие ресурсы, которые требуются для проведения технического обслуживания. Данный подход позволяет оценить, насколько техническое обслуживание является экономически выгодным, основываясь на сравнении выгод от обслуживания с затратами на его проведение.

Окупаемость технического обслуживания с применением предиктивной аналитики зависит от множества факторов, включая тип оборудования, стоимость обслуживания, вероятность отказов и потенциальные экономические потери от них. Окупаемость технического обслуживания по результатам предиктивной аналитики можно рассчитать с помощью формулы:

$$ROI = \frac{\Delta P}{C} \cdot 100\% , \text{ где} \quad (10)$$

ROI – индекс окупаемости,

ΔP – разница между доходом от использования предиктивной аналитики и затратами на её внедрение и обслуживание,

C – объем расходов на внедрение и обслуживание предиктивной аналитики.

При принятии решений на основе анализа данных с использованием алгоритмов машинного обучения и ИНС первым делом производится оценка качества модели с использованием метрик, и, если модель дает достаточно точные предсказания можно переходить к следующему шагу. По результатам предсказаний модели производится принятие решения о целесообразности проведения обслуживания. Например, если модель предсказывает высокую вероятность отказа оборудования или большие потери при отказе, будет целесообразно провести техническое обслуживание.

В исследованиях [147] приведен подход по принятию решений на основе данных классификации рабочего состояния агрегатных узлов для оптимизации расходов на обслуживание ветряных турбин.

$$TP_{savings} = \sum_{i=\#TP} (R - r) \left(\frac{\Delta t_i}{60} \right) \quad (11)$$

$$FN_{cost} = \#FN \times R^F FP_{cost} = \#FP \times I \quad (12)$$

$$Total_{savings} = \sum_{Turbine} \sum_{Component} (TP_{savings} - FN_{cost} - FP_{cost}), \text{ где} \quad (13)$$

$TP_{savings}$ – оценка экономии на основе обнаруженных сбоев,

FN_{cost} – оценка расходов стоимости замены оборудования на основе не выявленных сбоев,

FP_{cost} – оценка расходов стоимости проверки оборудования на основе ложных данных прогноза,

$Total_{savings}$ – итоговая оценка расходов стоимости технического обслуживания на основе анализа данных.

Данная математическая модель использует данные по классификации оценки текущего состояния ветряной турбины:

- Истинные положительные результаты (TP) – это сбои в работе ветряной турбины и подсистемы, обнаруженные за период от 2 до 60 дней до даты поломки. Если сбой обнаружен в нужное время, но в неправильной ветряной турбине или подсистеме, это засчитывается как ложноположительный результат. Истинные положительные стороны выражаются в экономии, которая представляет собой разницу между затратами на замену и ремонт.

- Ложноотрицательные сигналы (FN) – это реальные сбои в работе ветряной турбины и подсистемы, которые не были обнаружены в течение предыдущих 2-60 дней. Ложноотрицательные результаты переводятся в стоимость замены.

- Ложные срабатывания (FP) – это выданные предупреждения в ветряной турбине и подсистеме, в которых в течение следующих 260 дней не происходит сбоев. Ложные срабатывания переводятся в стоимость проверки.

Данный подход позволяет оценить затраты на ремонт оборудования и замену комплектующих в зависимости от технического состояния агрегатов ветряной турбины. После оценки расходов производится принятие решений по ремонту агрегата с последующей корректировкой плана технического обслуживания и назначением конкретных действий на операционном уровне.

2.6. Проектирование комплекса алгоритмов для проблемно-ориентированного конвейера обработки данных

В соответствии с результатами анализа подходов и методов принятий решений на основе данных в рамках предиктивной аналитики, сформированы следующие функциональные требования к алгоритмическому обеспечению:

1) Уровень данных:

- Экспорт и импорт данных («сырых», обучающих и тестовых датасетов) из систем хранения.
- Предварительная обработка данных (очистка, фильтрация выбросов, нормализация, преобразование, масштабирование, моделирование).
- Определение и отбор значимых признаков с использованием методов фильтрации и обертки.
- Разделение выборки на обучающий и тестовый датасеты.
- Поиск паттернов в датасетах с использованием методов оценки сходства данных.

2) Уровень моделей анализа данных:

- Доступ к разным источникам генерации моделей анализа данных посредством программных библиотек (таких как: Sklearn, Keras, TensorFlow).

- Поддержка различных алгоритмов машинного обучения, таких как классификация, регрессия и кластеризация.
- Автоматический подбор оптимальных гиперпараметров модели с помощью средств конвейера моделей и автоматического машинного обучения.
- Корректировка параметров модели в итеративном процессе обучения-тестирования.
- Предоставление метрик оценки качества моделей, таких как точность, полнота, F-мера, среднеквадратичная ошибка.
- Поддержка кросс-валидации для оценки обобщающей способности моделей.
- Поддержка онлайн и/или трансферного обучения, позволяющая обновлять модели по мере поступления новых данных.
- Возможность объединения нескольких моделей в ансамбли для повышения качества предсказаний с использованием методов бустинга и стекинга.
- Экспорт и импорт обученных моделей, показывающих удовлетворительные результаты тестирования.
- Оценка состояния и работоспособности оборудования.
- Подбор и рекомендация обученных моделей для анализа датасетов, имеющих паттерны данных с наибольшей степенью сходства.

3) Уровень принятия решений:

- Интеграция с инструментами визуализации для наглядного представления результатов анализа данных посредством программных библиотек (таких как: Seaborn, Matplotlib, Plotly).
- Экспорт и импорт данных в системе хранения регламента ТОиР, история операций, план ТОиР, стоимость компонентов и типов выполняемых работ.

- Выполнение расчета экономической эффективности мероприятий ТООР на основе метрик тестирования моделей анализа данных.

Исходя из выявленных ограничений к проектированию решения для выполнения задач в рамках диссертационного исследования предложена следующая алгоритмическая модель по подбору модели машинного обучения для анализа данных с парка разнородного оборудования (на примере промышленных роботов разных типов), которая включает в себя три основных блока:

- 1) Конструирование базовой модели на обучающей выборке:
 - 1.1) Получение массива исторических данных о работе промышленного робота; подготовка массива данных: нормализация, рандомизация, отбор значимых параметров, разделение на обучающую и тестовую выборку.
 - 1.2) Выбор метода и настройка параметров обучения модели.
 - 1.3) Тестирование метрик модели; валидация результатов прогнозирования для оценки решения поставленной задачи.
 - 1.4) На основе результатов пункта 3: настройка параметров и гиперпараметров модели; использование ансамблевых методов для увеличения точности прогнозирования.
- 2) Рекомендация к выбору обученной модели для тестовой выборки:
 - 2.1) Получение массива исторических данных о работе промышленного робота; подготовка массива данных: нормализация, рандомизация, отбор значимых параметров.
 - 2.2) Вычисление оценки сходства между тестовой выборкой и обучающей из пункта 1.1; определение порогового значения меры сходства.
 - 2.3) Тестирование базовой модели, обученной на выборке из пункта 1.1 на тестовой выборке, при прохождении определенного порога сходства.
 - 2.4) Оценка метрик и вывод о применимости модели для прогнозирования данных на тестовой выборке.
- 3) Принятие решений на основе данных

3.1) Оценка экономической эффективности для проведения ТОиР на основе данных (метрик тестирования модели).

3.2) Принятие решений о редактировании плана ТОиР и корректировке операционных действий.

3.3) Проведение операционных действий по ТОиР.

На рисунке 22 представлена диаграмма процесса конструирования базовой модели: от этапа управления данными по конкретному оборудованию до корректировки плана технического обслуживания.

На рисунке 23 представлена диаграмма подбора моделей на основе оценки подобия данных между обучающей выборкой (сформированной с одного промышленного робота) и тестовыми (сформированными с прочих промышленных роботов в рамках одного парка оборудования).

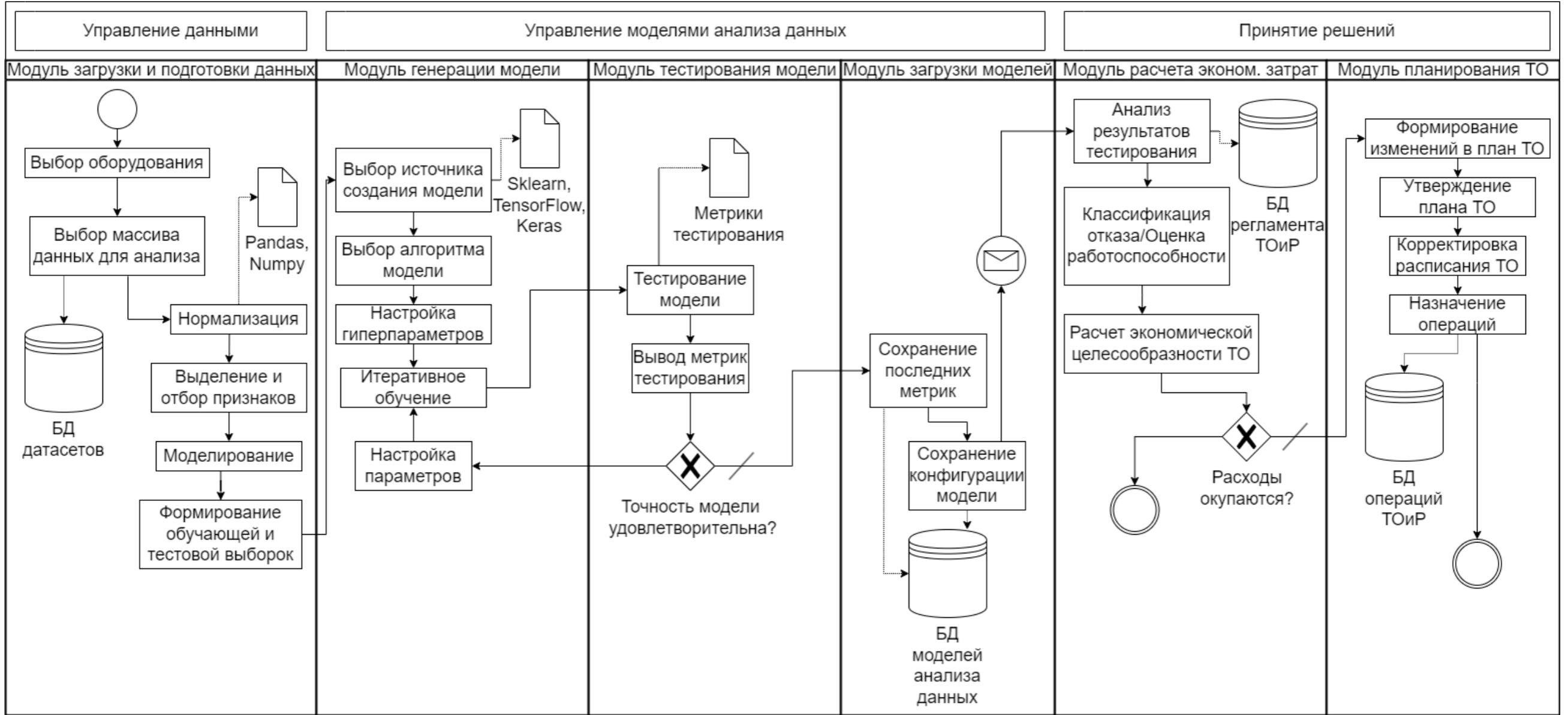


Рисунок 22 – Диаграмма процесса принятия решений ТОиР на основе конструирования базовой модели анализа данных

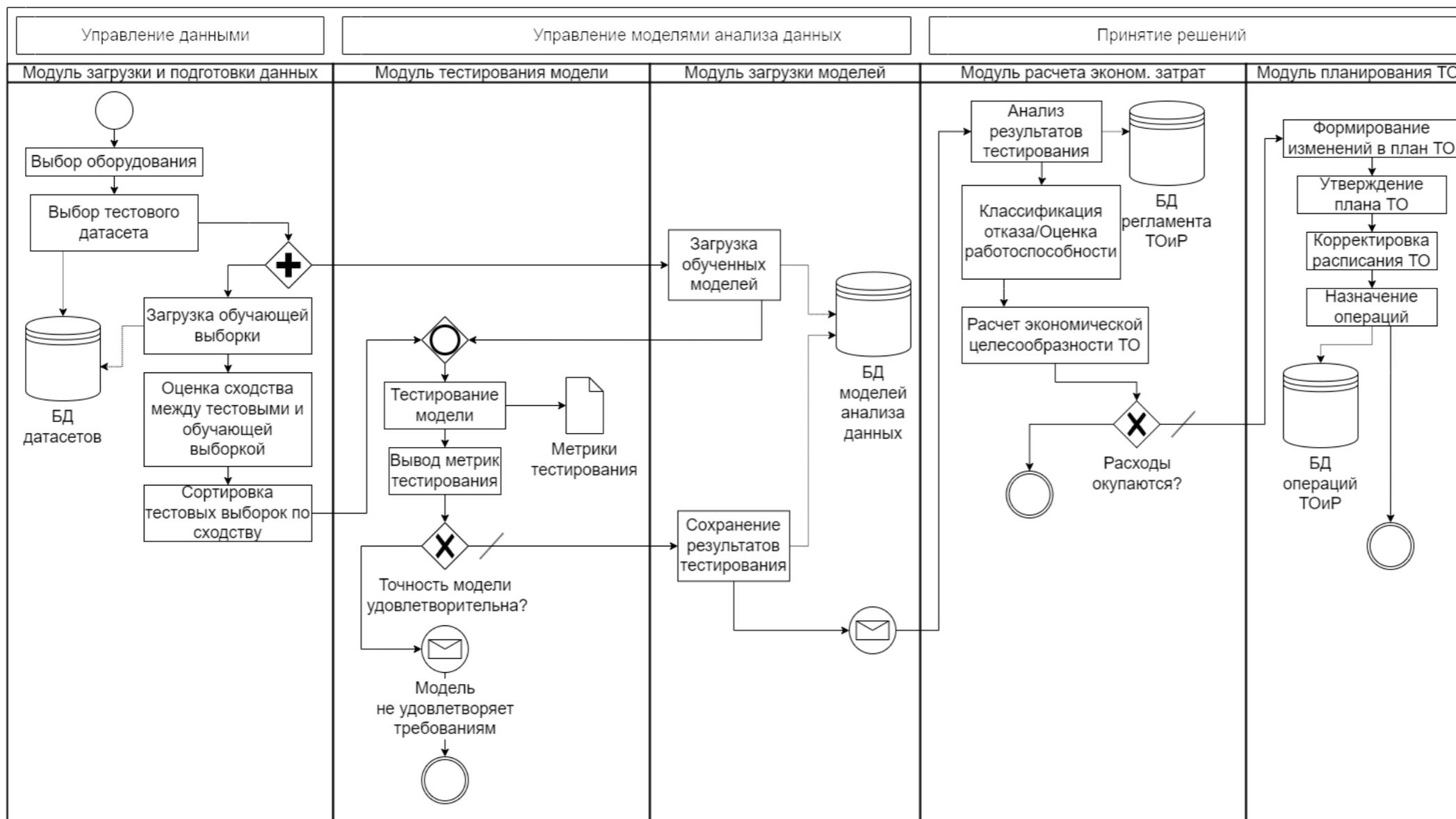


Рисунок 23 – Диаграмма подбора обученных моделей для тестовых выборок, собранных с промышленных роботов на основе оценки сходства данных

2.7. Выводы по ГЛАВЕ 2

Рассмотренные особенности процесса принятия решений по ТОиР позволили формализовать систему требований к проблемно-ориентированному конвейеру данных, а также реализовать отдельные его составляющие, а именно:

1) Комплекс моделей, обеспечивающих представление процесса принятия решений ТОиР в рамках предиктивного подхода на основе данных.

2) Архитектура программного обеспечения, реализующего построение проблемно-ориентированного конвейера для сбора, обработки и анализа данных промышленных роботов с целью принятия решений на основе данных для предиктивного технического обслуживания.

3) Модель принятия решений на основе данных, позволяющая произвести оценку экономической эффективности для принятия решений по ТОиР на основе данных.

4) Комплекс алгоритмов для проблемно-ориентированного конвейера, позволяющий реализовать процесс управления моделями анализа данных для адаптивного подбора и рекомендации модели на основе методов оценки подобия данных, собранных с парка промышленных роботов.

3. Программное обеспечение автоматизации процессов ТОиР, реализующего разработанный комплекс алгоритмов

3.1. Проектирование архитектуры программной системы

3.1.1. Описание архитектуры на основе выявленных ограничений

Исходя из ограничений и разработанной алгоритмической модели во второй главе, спроектирована следующая архитектура программной системы (рисунок 24). Система включает в себя следующие программные модули:

1) Модуль управления данными о работе оборудования, включающий в себя следующие функции:

- Импорт данных из хранилища исторических данных о работе оборудования.
- Агрегирование данных.
- Формирование выборок для анализа.
- Предобработка данных.
- Моделирование данных.
- Извлечение признаков.
- Поиск паттернов в выборках методами оценки сходства данных.
- Выявление аномальных значений.

2) Модуль анализа данных, включающий в себя следующие функции:

- Подготовка данных.
- Выбор источника генерации модели.
- Выбор алгоритма анализа данных.
- Подбор обученных моделей для анализа выборки.
- Настройка параметров и гиперпараметров модели.
- Отбор значимых признаков.
- Обучение и тестирование модели.
- Экспорт и импорт прогнозных моделей в хранилище.

3) Модуль оценки экономической эффективности, включающий в себя следующие функции:

- Оценка вероятности, временного периода до отказа оборудования.
- Оценка затрат на проведение ТОиР на основе анализа данных.
- Расчет издержек от простоя оборудования вследствие конкретного типа отказа.

4) Модуль поддержки принятия решений, включающий в себя следующие функции:

- Оценка экономических расчетов по результатам работы прогнозной модели.
- Планирование ТОиР и рекомендации о корректировке плана с учетом экономической целесообразности.
- Валидация модели анализа данных.

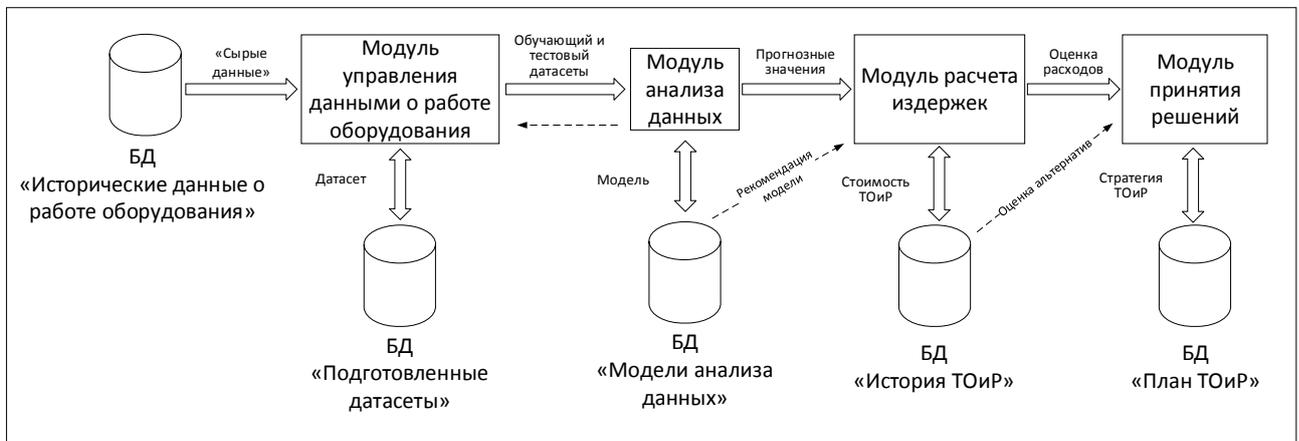


Рисунок 24 – Концептуальная схема программной архитектуры

Декомпозиция процесса планирования ТОиР применима для анализа данных в условиях обучения модели (или нескольких моделей) анализа для каждой производственной единицы оборудования в рамках подхода трансферного обучения. Основываясь на разработанной алгоритмической модели из главы 2, производится корректировка процесса А2 для использования обученных моделей на тестовых выборках, сформированных на прочих производственных единицах

(рисунок 25). Процесс включает в себя алгоритм поиска паттернов в датасетах с использованием методов оценки сходства данных.

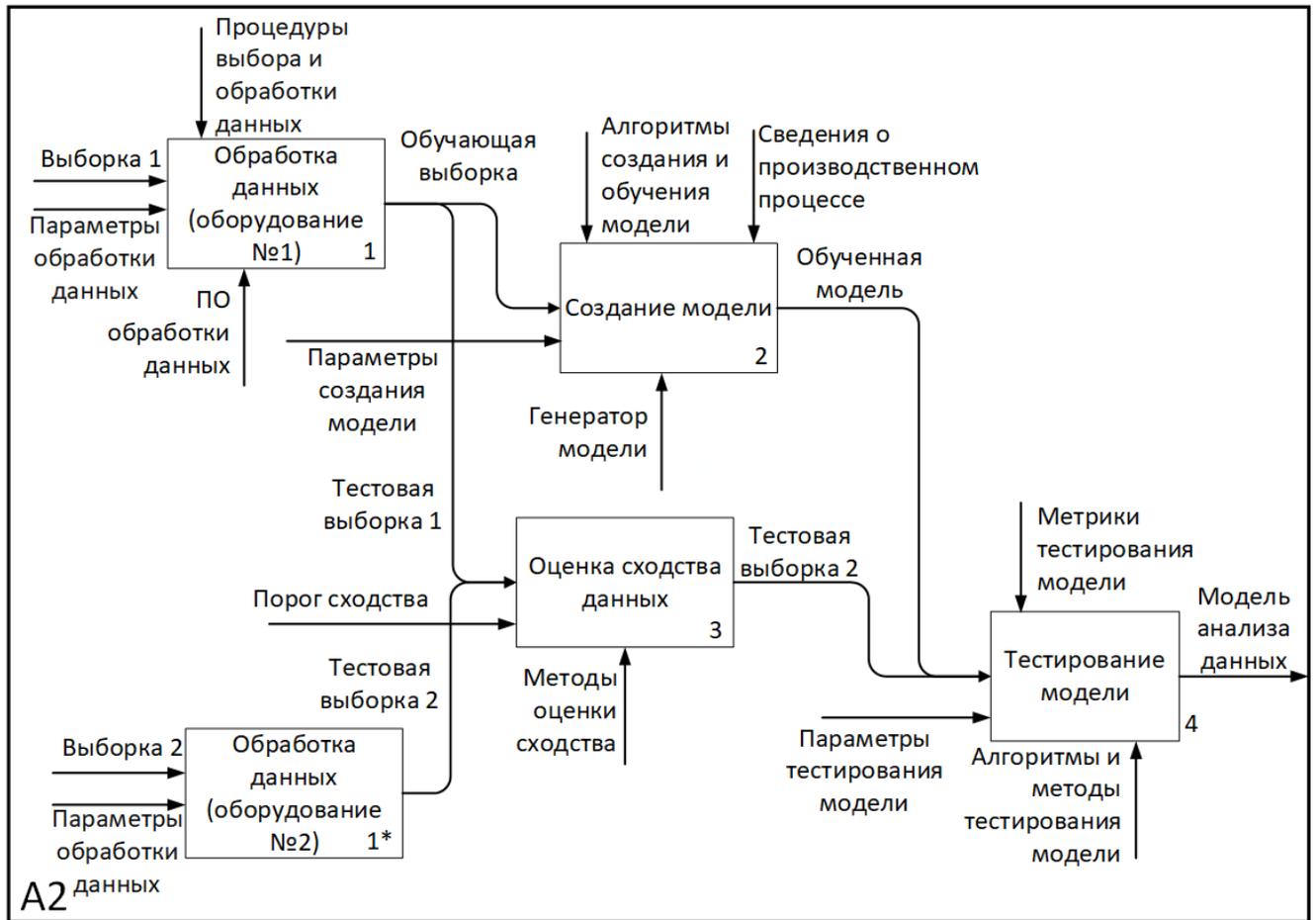


Рисунок 25 – Процесс тестирования моделей на тестовых выборках с определенной степенью сходства (A2)

3.1.2. Диаграмма компонентов программного обеспечения

На основе состава программной системы (пункт 3.1.1) сформирована следующая диаграмма компонентов с использованием нотации UML (рисунок 26). Диаграмма компонентов определяет важность для проектирования системы принятия решений на основе анализа данных по ряду следующих причин:

1) Диаграмма позволяет визуально представить структуру системы и ее компоненты, что позволяет наглядно представить взаимодействие и типы связей между элементами системы.

2) Данная нотация определяет компоненты, их функции и связи между ними, что важно для разработки архитектуры системы на основе требуемых ограничений.

3) При проектировании компонентов предопределяется возможность учесть масштабируемость системы и разделить ее на модули для удобства управления и разработки программного обеспечения.

Сформированная диаграмма компонентов наглядно представляет состав разработанной программной системы, что позволяет оценить полноту соблюдения требований. Одной из ключевых особенностей можно выделить возможность отображения подключаемых компонентов «извне» (то есть: программные библиотеки, подключенные через систему управления пакетами). Перечень используемых программных библиотек:

- Pandas – библиотека для работы с табличными данными и временными рядами.
- Numpy – библиотека для работы с многомерными массивами и математическими функциями.
- Statistics – библиотека, предоставляющая статистические функции.
- Django Framework – высокоуровневый фреймворк, позволяющий создавать веб-сайты.
- Matplotlib, Plotly – библиотеки визуализации для создания графиков и диаграмм.
- Sklearn, Keras, Tensorflow – библиотеки для работы с алгоритмами машинного обучения и ИНС.
- PyMongo – библиотека-провайдер для связи с нереляционной базой данных MongoDB.

- Psycopg2 – библиотека-провайдер для связи с реляционной базой данных PostgreSQL.

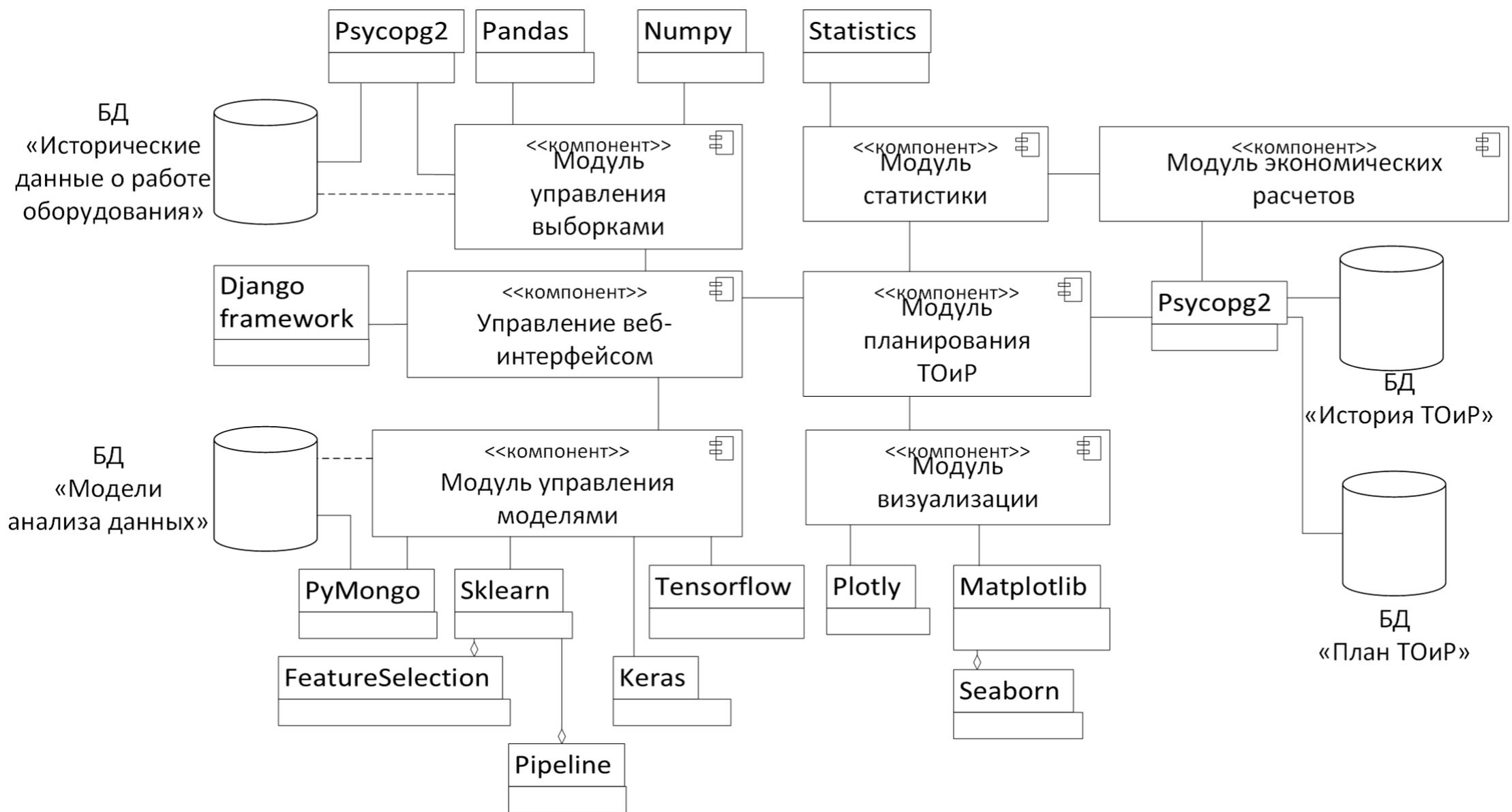


Рисунок 26 – Диаграмма компонентов программной системы

3.2. Модуль управления данными

3.2.1. Диаграмма вариантов использования модуля управления данными

В соответствии с описанным перечнем функций разработана следующая диаграмма вариантов использования для пользователя (рисунок 27). В качестве пользователя модуля управления данными выступает аналитик данных, совершающий агрегирование, подготовку и моделирование данных с последующей передачей на этап создания моделей и анализа.



Рисунок 27 – Диаграмма вариантов использования модуля управления данными

Диаграмма вариантов использования позволяет отобразить основные требования к разрабатываемой системе, на которой отражаются границы системы (актеры) и ее основная функциональность (прецеденты).

3.2.2. База данных хранения исторических данных о работе оборудования (на примере промышленных роботов)

Для базы данных хранения параметров мониторинга промышленных роботов была разработана следующая структура (логическая схема БД представлена на рисунке 28).

База данных содержит следующие таблицы:

- 1) `Token` – создает токен доступа для установки рабочей программы промышленного робота.
- 2) `User` – пользователь системы мониторинга и установки рабочих программ.
- 3) `UserRole` – роль пользователя для авторизации.
- 4) `Robot` – параметры настройки промышленного робота.
- 5) `RobotRole` – название исполняемой роли промышленного робота.
- 6) `WorkingProgram` – рабочая программа (состоящая из последовательности команд) для выбранного промышленного робота.
- 7) `ApplicationArea` – участок/траектория/путь для работы промышленного робота.
- 8) `Parameter` – массив параметров, сохраняемых в таблицу, в процессе выполнения рабочей программы промышленного робота.

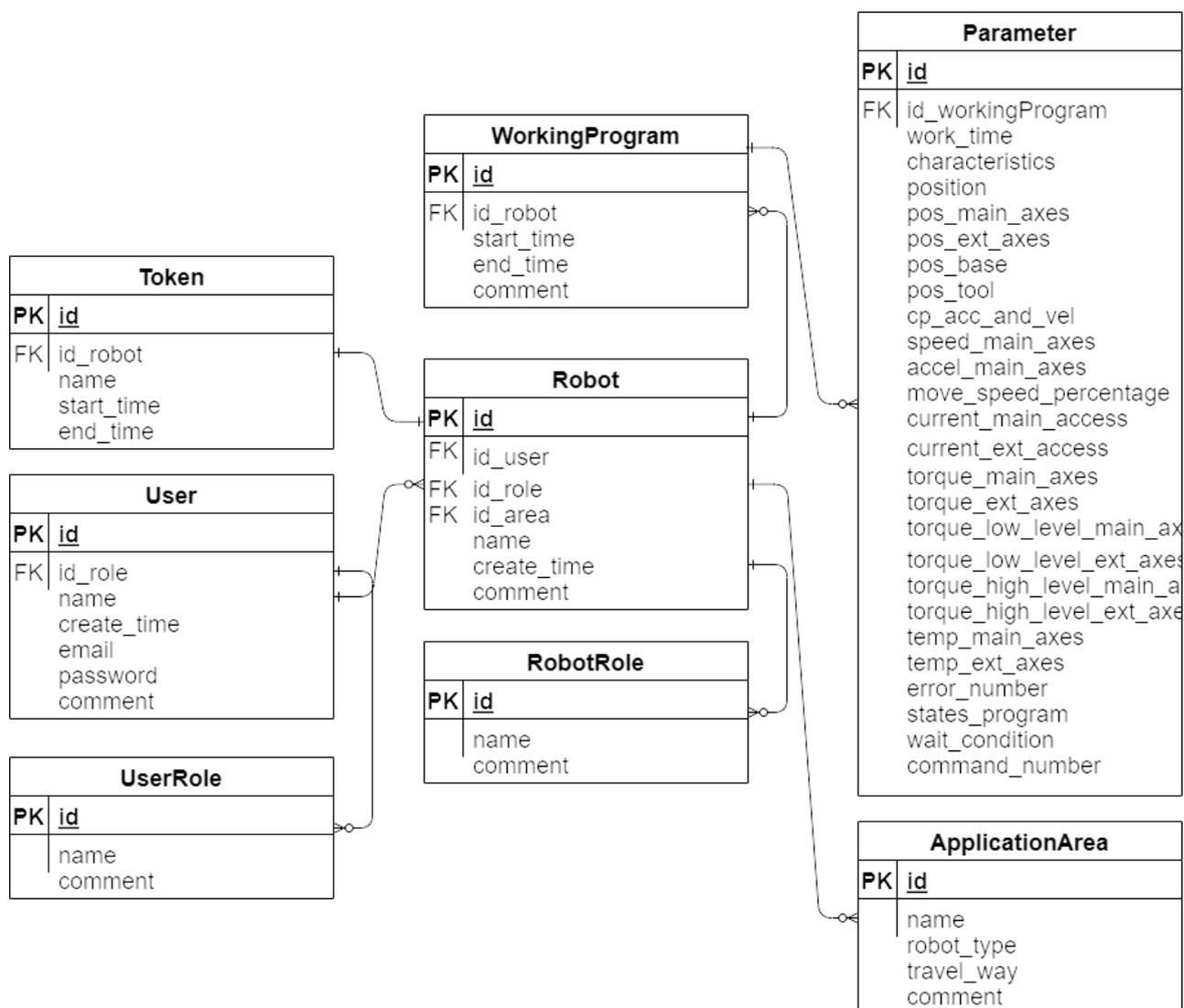


Рисунок 28 – Схема базы данных хранения массивов исторических данных о работе промышленных роботов

Алгоритм для процесса загрузки и формирования выборки для анализа представлена на рисунке 29. Пользователь, используя форму загрузки данных в интерфейсе загружает новый массив данных. Загруженные файлы передаются на исполнение в функцию обработки данных и сохраняются в оперативной памяти в качестве выборки.

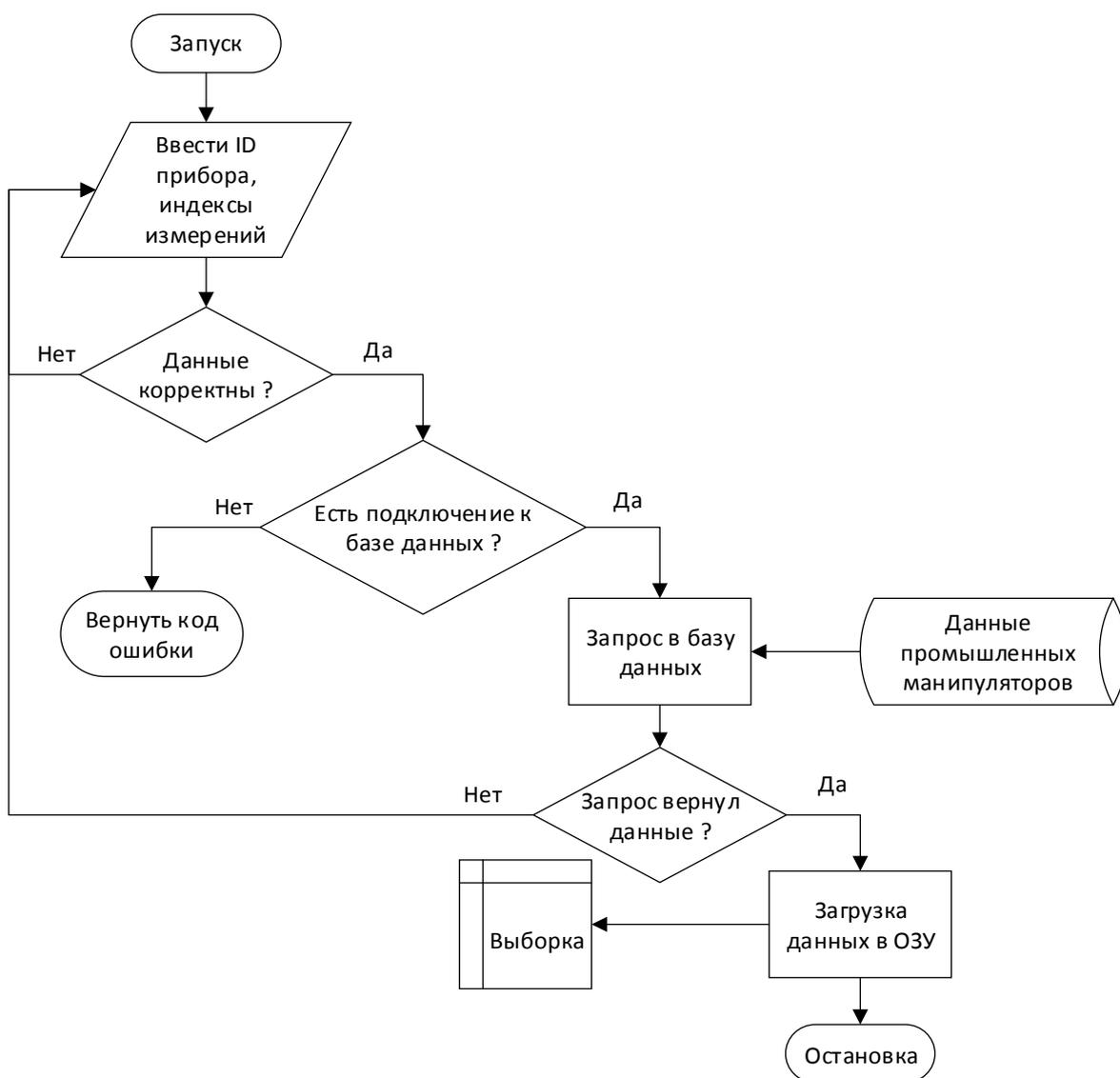


Рисунок 29 – Алгоритм загрузки выборки из базы данных

3.2.3. Модуль выявления аномальных значений

Аномальные значения при анализе больших данных представляют собой отклонения от общего тренда или нормы в наборе данных. Они могут быть вызваны ошибками в сборе данных, случайными факторами или наличием некорректных значений. Для обнаружения аномалий используются статистические методы, машинное обучение или визуализация данных.

Аномалии могут быть типизированы следующим образом:

- Выброс – значение, значительно удаленное от основного тренда или кластера распределения данных; алгоритм выявления выброса интуитивно понятен

– необходимо определить, насколько далеко располагается отдельное значение от среднего. Основные методы поиска выбросов: правило трех сигм, Z-оценка, межквартильный размах, тест Граббса и др.

- Сдвиг – проявляется внезапным изменением показателей или распределения данных относительно предыдущих значений без явной причины или объяснения. Основные методы поиска сдвигов: контрольные карты Шухарта, методы машинного обучения (k -средние, одноклассовая классификация), автоэнкодер и др.

- Отклонение от тренда (для временных рядов с фактором сезонности) – для обнаружения аномалии необходимо сравнить текущий период и несколько предыдущих. Одним из эффективных подходов является метод Хольта-Винтерса.

В исследовании [166] рассмотрен алгоритм поиска аномальных значений в контексте управления процессом мониторинга производственного процесса промышленного манипулятора. Под аномальными значениями авторы подразумевают параметры работы манипулятора (такие как: момент силы, ток, температура привода), которые выходят за границы средних значений для каждой из рабочих операций, выполняемых роботом по изначально заданной оператором программе. Предложенный алгоритм является улучшением штатного алгоритма обнаружения аномалий, который не предполагает гибкой настройки диапазона допустимых значений при настройках на определенный перечень рабочих операций (алгоритм представлен в приложении 11).

$$\bar{I}_{O^k, M_x}^n = \frac{\sum_{i=O_{t_1}^k}^{O_{t_2}^k} (I_{i, M_x}^n)}{O_{t_2}^k - O_{t_1}^k}, \text{ где} \quad (14)$$

n – номер оси (аналогичен номеру электропривода),

O^k – операция под номером k ,

\bar{I}_{O^k, M_x}^n – значение тока для оси n , в момент времени i , при перемещении груза

M_x .

Для получения коридора возможных значений (верхняя и нижняя граница значений токов при работе робота с грузами разной массы) необходимо получить перечень максимальных и минимальных средних значений тока по операциям:

$$\text{Max} \left(\bar{I}_{O_j^k, M_x}^n \right), \quad (15)$$

$$\text{Min} \left(\bar{I}_{O_j^k, M_x}^n \right), \text{ где} \quad (16)$$

j – номер итерации при выполнении операции O^k с грузом M_x .

Далее необходимо получить максимальные и минимальные значения токов для операций со всеми значениями массы дополнительного груза:

$$N_{\text{max}_{O^k}} = \text{Max} \left(\text{Max} \left(\bar{I}_{O_j^k, M_x}^n \right) \right), \quad (17)$$

$$N_{\text{min}_{O^k}} = \text{Min} \left(\text{Min} \left(\bar{I}_{O_j^k, M_x}^n \right) \right), \text{ где} \quad (18)$$

$N_{\text{max}_{O^k}}$ – максимально допустимое значение тока во время выполнения операции O^k ;

$N_{\text{min}_{O^k}}$ – минимально допустимое значение тока во время выполнения операции O^k .

Данная операция позволяет построить верхнюю и нижнюю границу диапазона допустимых значений при всех циклах работы манипулятора.

В исследовании [167] предлагается улучшение для выше обозначенного алгоритма посредством применения метода аппроксимации (для сглаживания значений электрического тока) и классификации нагрузки на манипуляторе с помощью следующих методов: дерево решений, наивный Байес, метод опорных векторов (SVM). Применение метода аппроксимации позволяет улучшить точность классификации за счет сглаживания и нивелирования лишнего «шума» в наборах данных. Применение такого рода алгоритмов позволяет не только детектировать перегрузку на манипуляторе, но и производить настройку/мониторинг

производственного процесса со своевременными уведомлениями об изменениях в измеряемых параметрах манипулятора, выходящих за границы настроек технологического процесса выполняемой оборудованием программой.

3.2.4. Алгоритм оценки сходства данных

Для выбора наиболее подходящего датасета тестирования базовой модели производится оценка сходства выборок, сформированных с данных мониторинга промышленного оборудования. Тестовые выборки и обучающий датасет для базовой модели хранятся в базе данных. Результатом работы алгоритма является перечень коэффициентов сходства между обучающей и тестовой выборкой для каждой из тестовых выборок. Блок-схема алгоритма представлена на рисунке 30.

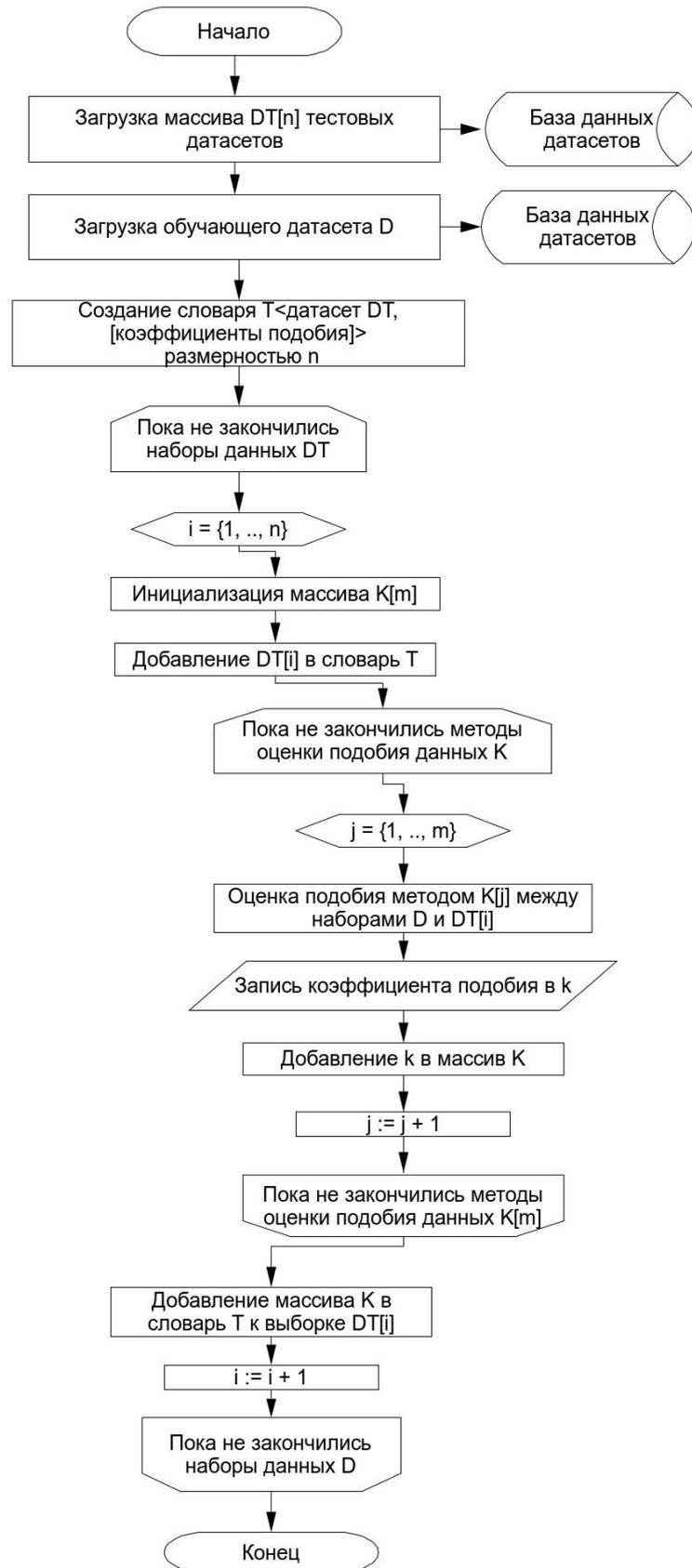


Рисунок 30 – Алгоритм оценки сходства датасетов

3.3. Модуль управления моделями анализа данных

3.3.1. Диаграмма вариантов использования модуля анализа данных

Диаграмма вариантов использования представлена на рисунке 31. В качестве пользователя модуля анализа данных выступает аналитик данных. Основной функцией данного модуля является реализация процесса управления интеллектуальными моделями анализа данных. Исходя из требований процесса управления выделяются функции по созданию, настройке, подборке, тестирования и валидации модели для анализа выбранных датасетов.



Рисунок 31 – Диаграмма вариантов использования модуля анализа данных

3.3.2. Алгоритм обучения и тестирования модели анализа данных

В общем виде алгоритмы обучения и тестирования моделей интеллектуального анализа данных представлены в приложении 6. Для программной реализации спроектирована блок-схема алгоритмов тестирования моделей на данных, собранных с разных единиц оборудования (рисунок 32).

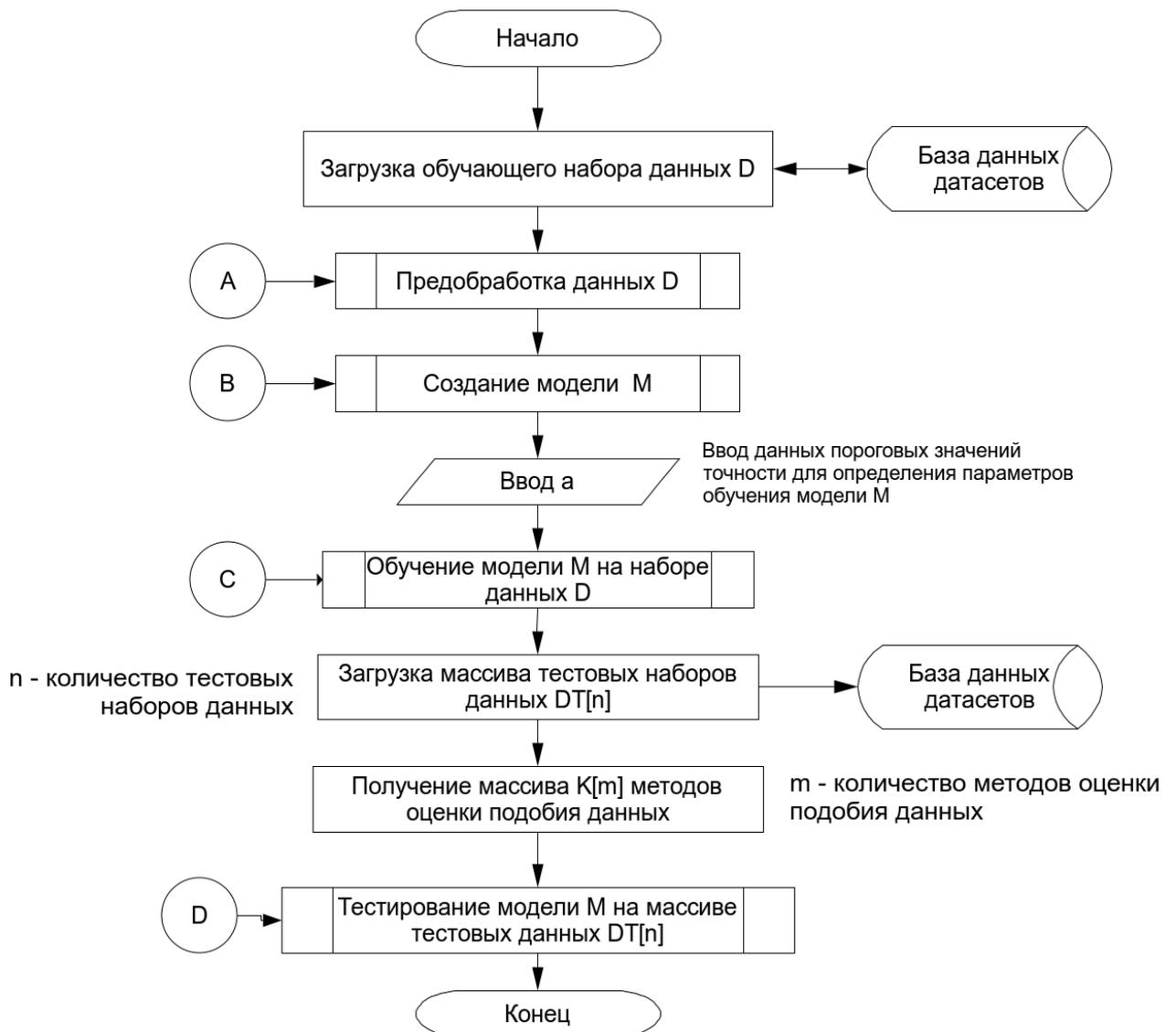


Рисунок 32 – Алгоритм обучения и тестирования модели на наборах данных разнородного оборудования

Более подробная детализация функциональных блоков (под ключевыми идентификаторами А, В, С, D) алгоритма представлена в приложении 5.

3.3.3. Алгоритм подбора модели машинного обучения на основе оценки сходства данных

На этапе решения задачи подбора моделей машинного обучения на основе оценки сходства данных предложено производить кластеризацию экспериментальных данных с использованием алгоритма k -средних, в качестве расстояния между объектами кластера используется мера сходства наборов данных. Предлагаемый алгоритм состоит из следующих основных шагов:

1) Вычисление сходства временных рядов – производится оценка меры сходства каждого входного набора данных с каждым другим одним из предложенных ранее методов (Евклидово расстояние, алгоритм трансформации временной шкалы и т.д.).

2) Кластеризация наборов данных – с использованием метода k -средних формируются множества кластеров, в качестве расстояния между объектами используется мера сходства наборов данных.

3) Поиск ближайшего к центроиду набора данных – в сформированных кластерах осуществляется поиск ближайшего к центроиду набора данных, на основе которого строится эталонная для данного кластера модель машинного обучения.

4) Построение модели машинного обучения – выбранные наборы данных используются для построения моделей машинного обучения и ИНС.

Результатом данного этапа является набор сформированных «эталонных» модели машинного обучения и соответствующие им наборы данных наиболее близких к центроидам кластеров. На рисунке 33 представлен пример кластеризации нескольких наборов данных.

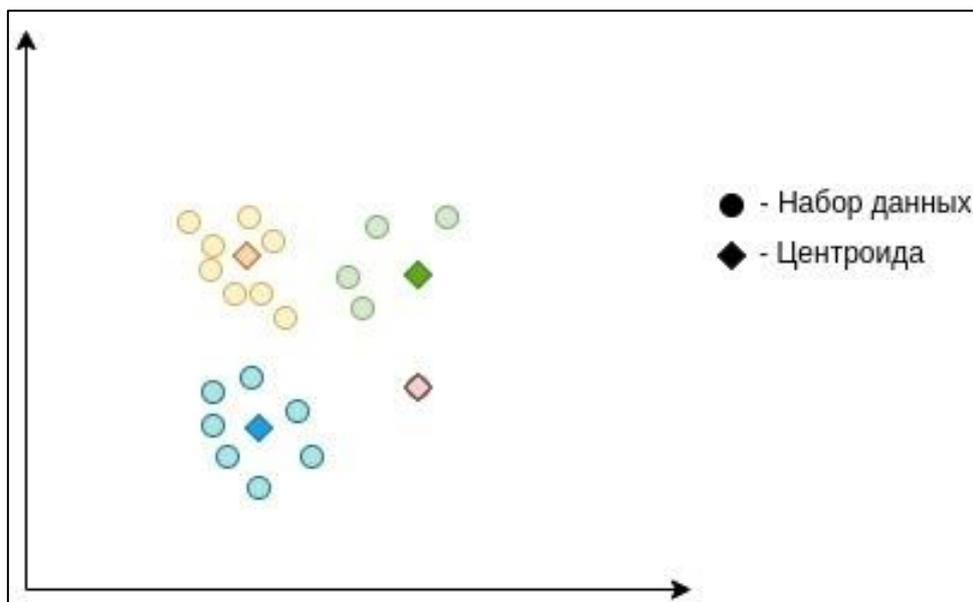


Рисунок 33 – Кластеризация наборов данных методом k -средних

Последним шагом этапа инициализации производится проверка выбранных эталонных моделей на тестовых наборах данных, принадлежащих кластеру. В качестве проверки корректности разбиения наборов на кластеры строятся модели машинного обучения для прочих наборов данных, сравнивается точность модели обучения на основе основного набора и на основе набора данных ближайшего к центроиду кластера. Если полученная ошибка больше 50%, то происходит запуск алгоритма кластеризации с другими параметрами. На рисунках 34–36 представлена блок-схема описанных шагов этапа инициализации.

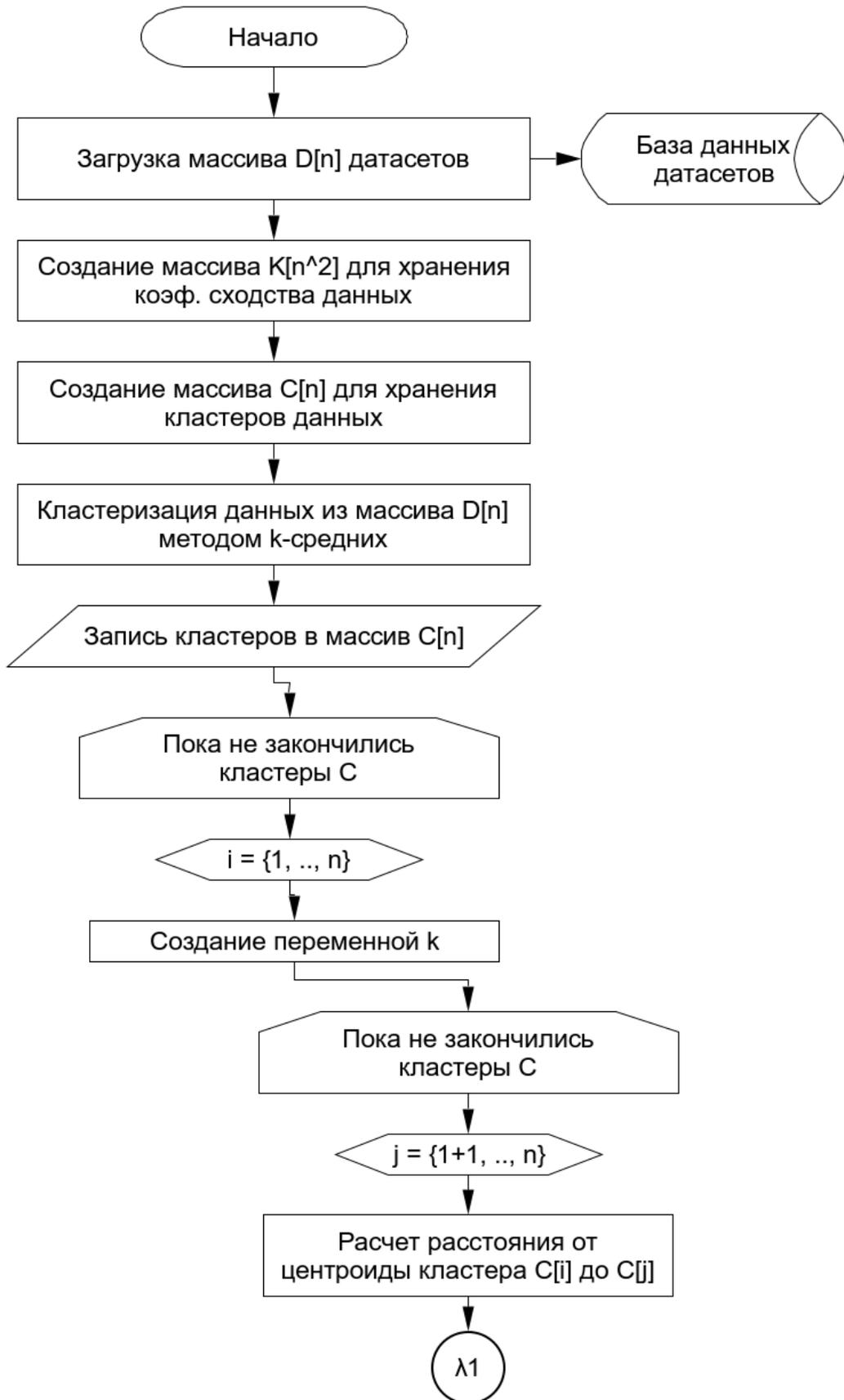


Рисунок 34 – Инициализация базовых моделей

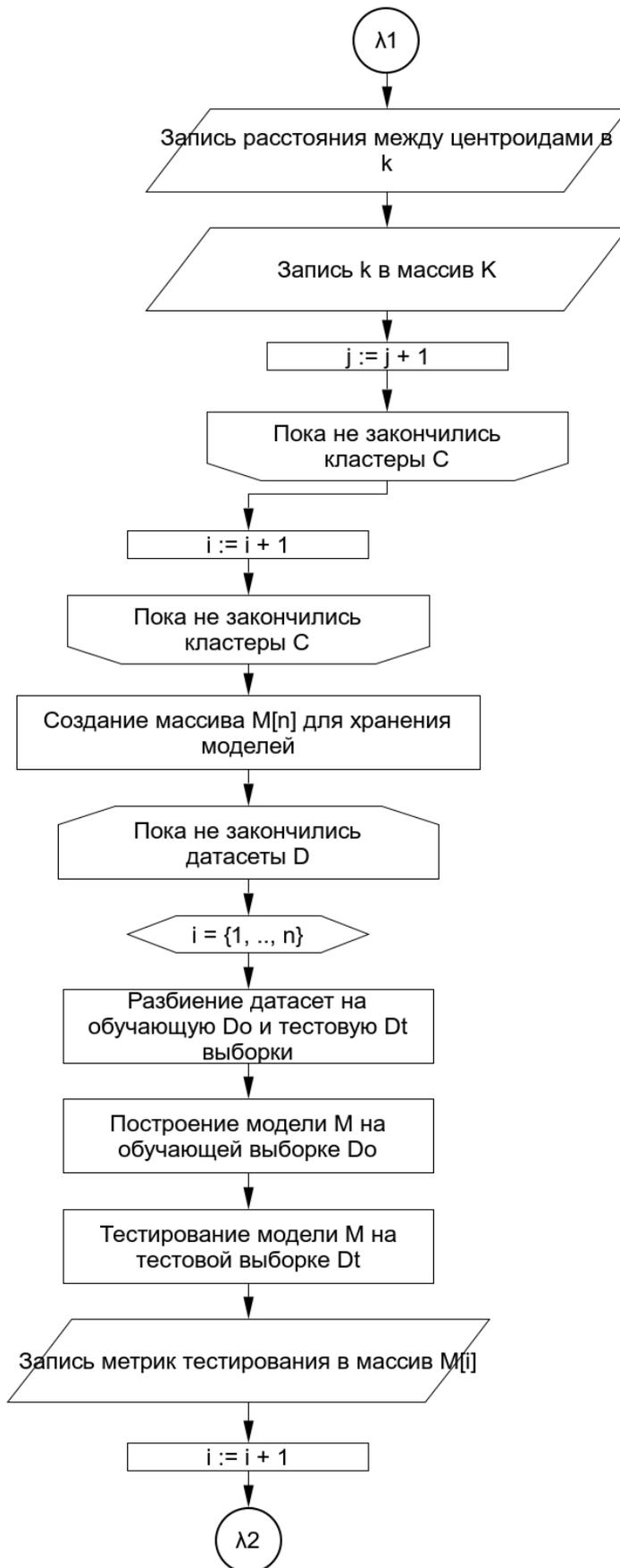


Рисунок 35 – Инициализация базовых моделей

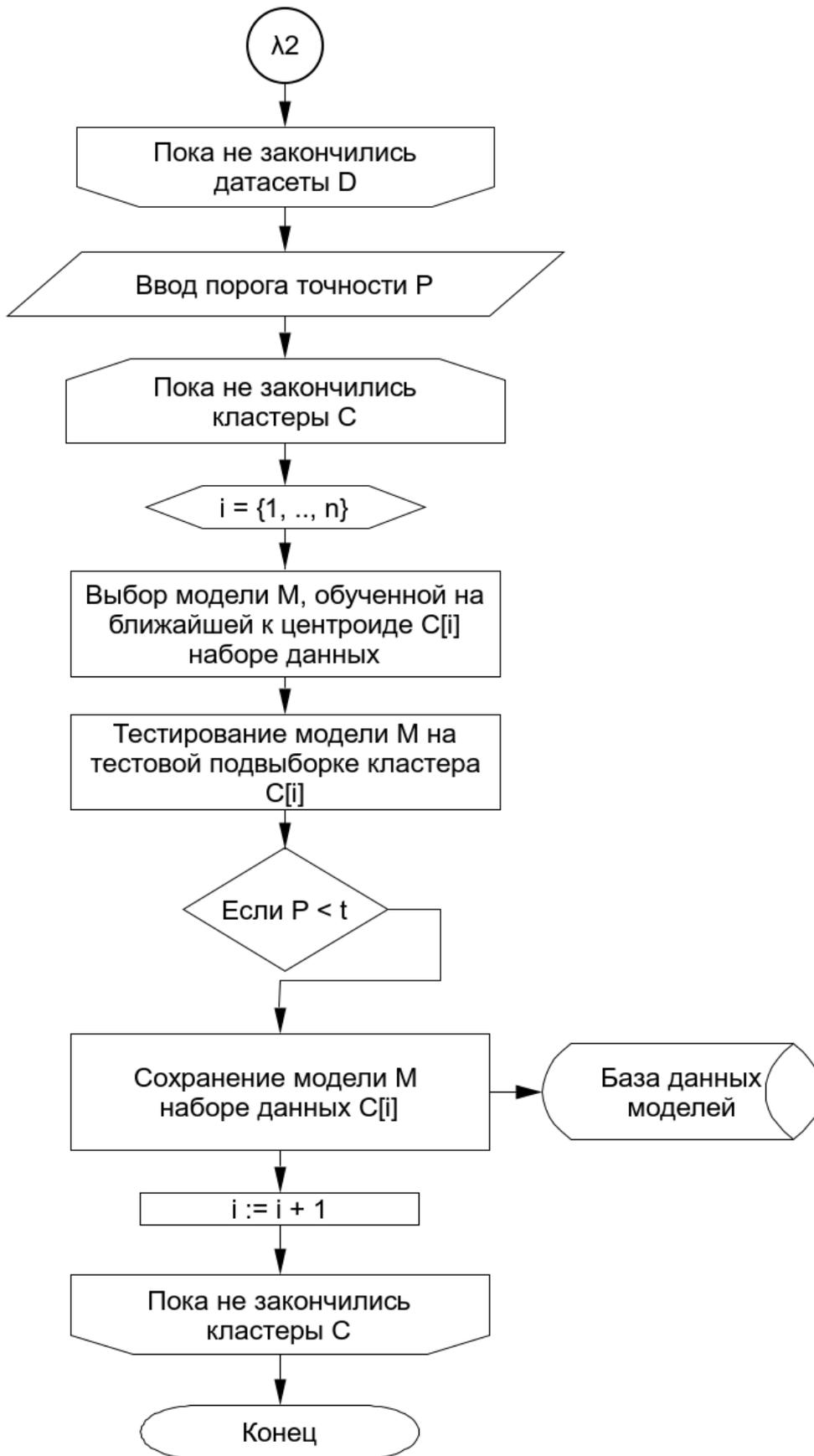


Рисунок 36 – Инициализация базовых моделей

На втором этапе набор эталонных моделей и соответствующие им выборки используются для наборов данных, имеющих наибольшую степень сходства. Предлагаемый алгоритм состоит из следующих шагов:

- 1) Вычисление сходства временных рядов – оценка меры сходства входного набора данных с набором эталонных моделей.
- 2) Выбор кластера – оценка принадлежности входного набора данных к имеющимся кластерам, выбор кластера с наименьшим расстоянием между входным набором и центроидом.
- 3) Использование модели машинного обучения – исходная выборка подается в качестве входного параметра для выбранной эталонной модели машинного обучения.

Для оценки эффективности подобранной модели были выбраны 2 метода:

- 1) MAE – метод средней ошибки. Выходное значение представляет собой процент ошибок от фактических значений. Использование метода абсолютной ошибки имеет смысл в контексте сравнения качества нескольких моделей, однако этот метод не позволяет определить качество модели.
- 2) MSE – метод среднеквадратичной ошибки. Данный метод обладает такими же свойствами, как метод абсолютной ошибки, однако более чувствителен к выбросам.

3.3.4. Диаграммы последовательности модуля анализа данных

На рисунке 37 представлена диаграмма последовательности алгоритма оценки точности подобранной модели машинного обучения в нотации UML.

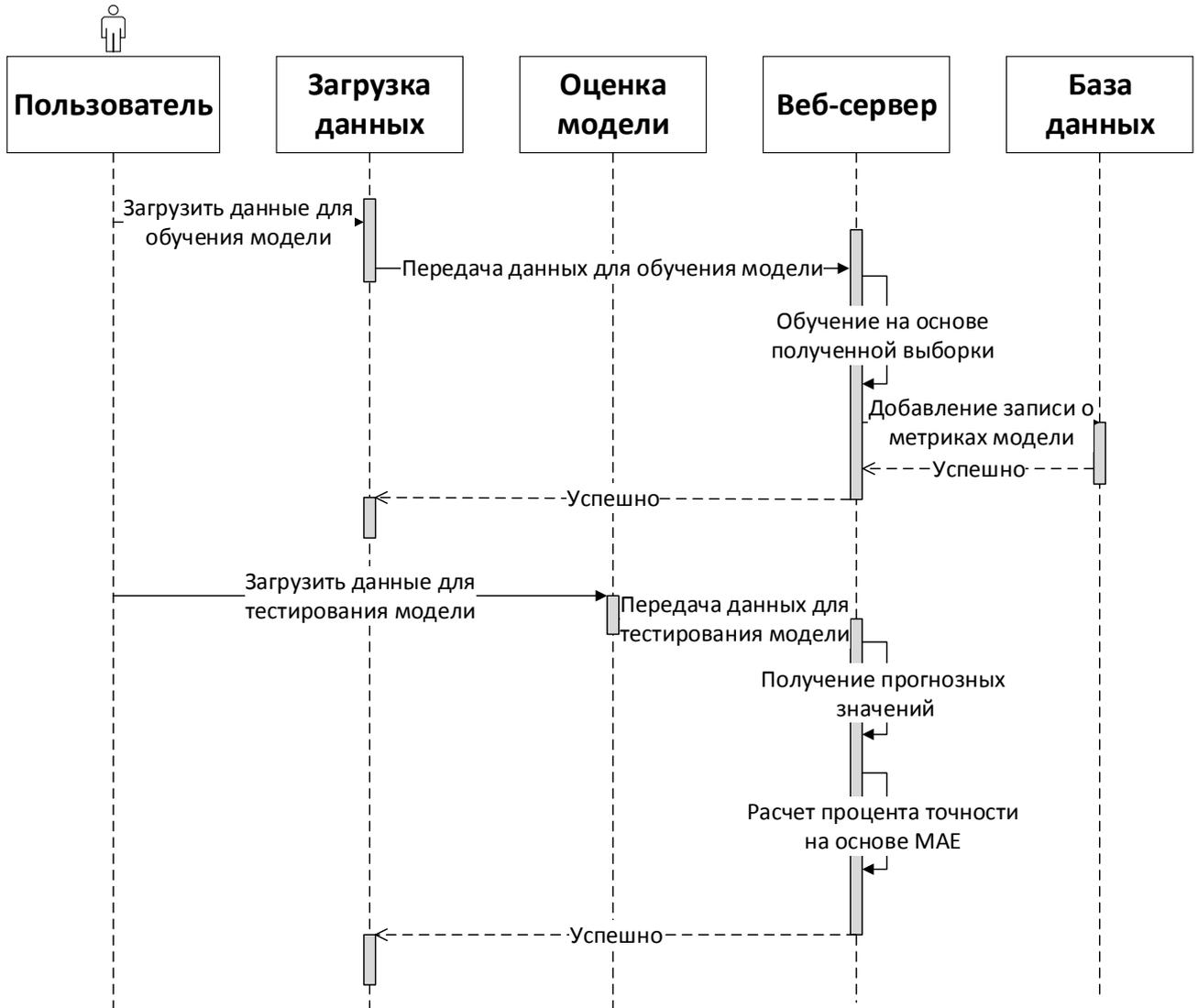


Рисунок 37 – Диаграмма последовательности алгоритма оценки модели

Алгоритм состоит из двух этапов: подготовка модели для оценки (обучение модели) и тестирование, в рамках которого будет произведена оценка точности модели.

Первый этап: обучение модели. На данном этапе строится модель машинного обучения на основе входного набора данных. Этап обучения модели состоит из следующих шагов:

1) Преобразование выборки – исходный набор данных разделяется в соотношении 70 на 30. Первая часть используется для обучения, вторая часть для тестирования модели.

2) Построение модели машинного обучения – полученная выборка используется для обучения модели: линейная регрессия, многослойный персептрон и рекуррентная ИНС (RNN).

3) Оценка ошибки – оценка полученной модели (метод абсолютной ошибки, метод среднеквадратичной ошибки).

Второй этап: тестирование. На данном этапе происходит сравнение ошибки, рассчитанной на первом этапе и ошибки подобранной модели. Если величина ошибки не превышает установленного значения, то алгоритм подбора модели машинного обучения на основе подобию исходных данных работает корректно.

Диаграмма последовательности для процесса обучения и тестирования модели анализа данных представлена на рисунке 38. Пользователь, используя интерфейс, загружает данные (из файла или готовую выборку из базы данных), далее выборка подготавливается для передачи в созданную модель анализа данных (производится разбиение на обучающую и тестовую выборку в соотношении 70:30). При успешном обучении модели производится загрузка в базу данных хранения моделей и передача модели на этап тестирования.

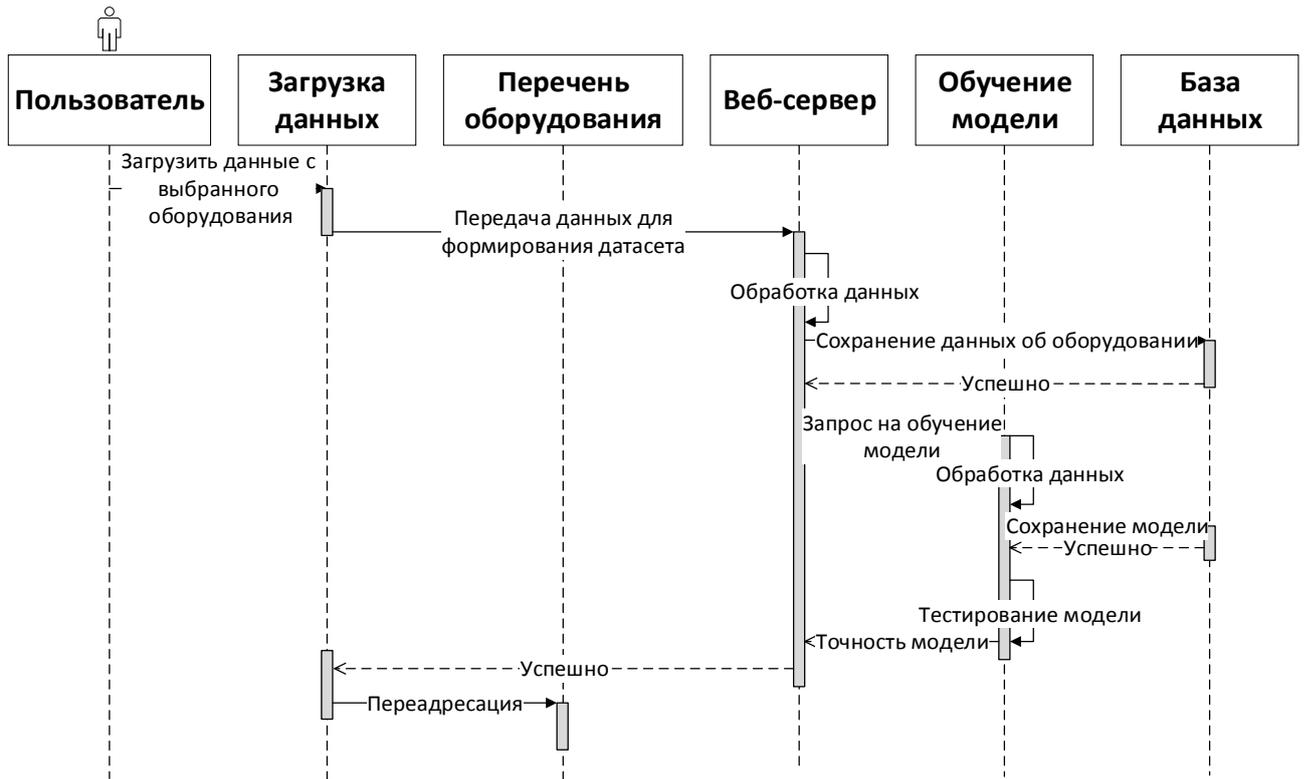


Рисунок 38 – Обучение и тестирование базовой модели

Диаграмма последовательности для алгоритма оценки точности модели на данных с определенным коэффициентом сходства представлена на рисунке 39. Пользователь выбирает промышленного робота и датасет, которому необходимо подобрать модель анализа. После выбора датасета производится запрос в базу данных хранения моделей для получения определенного списка моделей. Далее производится оценка сходства между выбранным датасетом и обучающими датасетами загруженных моделей. Далее производится тестирование двух моделей (датасеты которых имеют наибольшую степень сходства) и информация о тестировании выводится в интерфейс программного обеспечения.



Рисунок 39 – Оценка сходства данных и тестирование модели

Диаграмма последовательности для процесса обратной связи и визуализации результатов тестирования модели анализа данных представлена на рисунке 40. Визуализация позволяет наглядно продемонстрировать результаты в удобном для человека виде на интерактивной веб-форме либо в сформированной графическом файле в расширении PNG.

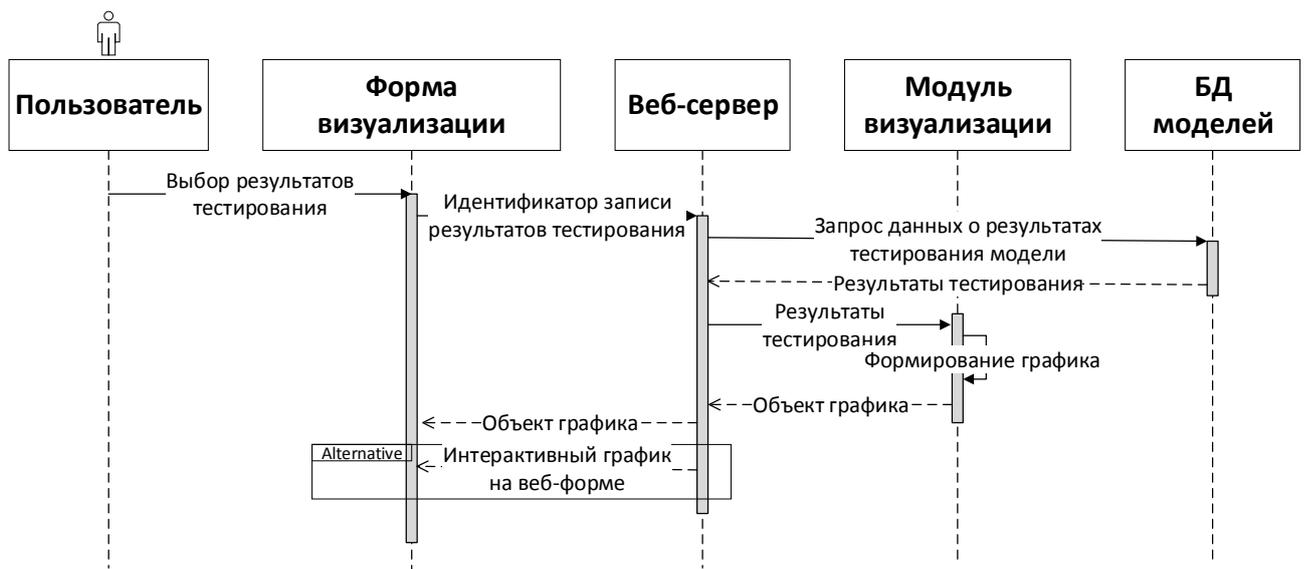


Рисунок 40 – Обратная связь и визуализация результатов тестирования

3.3.5. База данных для хранения моделей интеллектуального анализа

Метаданные моделей анализа данных и метрики тестирования хранятся в нереляционном хранилище (используемая СУБД для хранения данных – MongoDB). Схема базы данных учитывает выявленные ограничения и требования нормализации. Схема состоит из следующих сущностей (на прикладном уровне называющихся «коллекциями»):

- 1) Model – коллекция для хранения данных описания обученной и сохраненной модели.
- 2) ModelState – коллекция для хранения данных состояний сохраненных моделей.
- 3) ModelMetrics – коллекция для хранения массива метрик и истории тестирования моделей с указанием конкретных обучающих и тестовых выборок.
- 4) Dataset – коллекция для хранения обучающих и тестирующих датасетов, сформированных на данных мониторинга рабочих процессов промышленного оборудования (раздел 3.2.2).

На рисунке 41 представлена логическая схема БД с указанием конкретных атрибутов. Использование нереляционной базы данных для работы с моделями анализа данных и датасетами (сформированных с парка разнородного оборудования) имеет следующие преимущества:

- 1) Информация сохраняется в формате документов, что делает систему более гибкой и масштабируемой.
- 2) Работа с различными типами данных. Эта особенность имеет значение при обработке больших данных, не встраиваемых в одну структуру и собираемых из разных источников.
- 3) СУБД позволяет искать информацию по полям. Поддерживаются диапазонные запросы и поиск по регулярным выражениям.

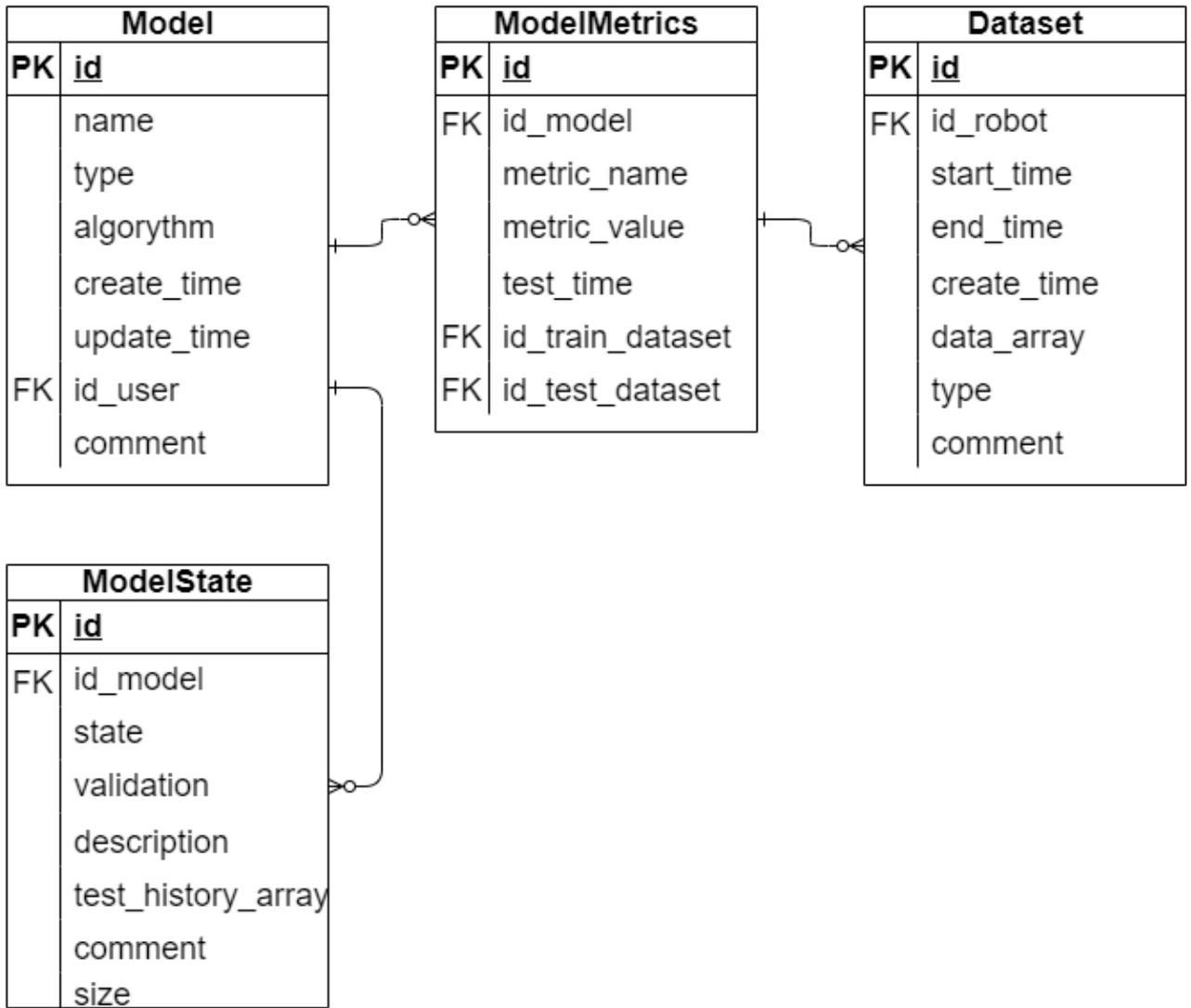


Рисунок 41 – База данных для хранения моделей и датасетов

3.4. Модуль поддержки принятия решений на основе анализа данных

Для хранения исторических данных о техническом обслуживании и ремонте сформирована следующая база данных (Приложение 4). Ключевыми сущностями данной БД являются:

- 1) Ремонтные работы (включают в себя: тип, время, категория, стоимость и расход материалов).
- 2) Работы по техническому обслуживанию (включают в себя: объект обслуживания, необходимое снаряжение, время, периодичность, стоимость, задействованные рабочие и их квалификация).

Описание процесса оценки расходов ТОиР на основе данных представлен на рисунке 42. Для принятия решений на основе анализа данных с использованием моделей анализа производится оценка качества модели с использованием метрик для выбранных данных, и, если модель дает достаточно точные предсказания, можно переходить к оценкам экономической выгоды альтернативных решений по ТОиР. По результатам прогнозных значений модели производится принятие решения о целесообразности проведения обслуживания.

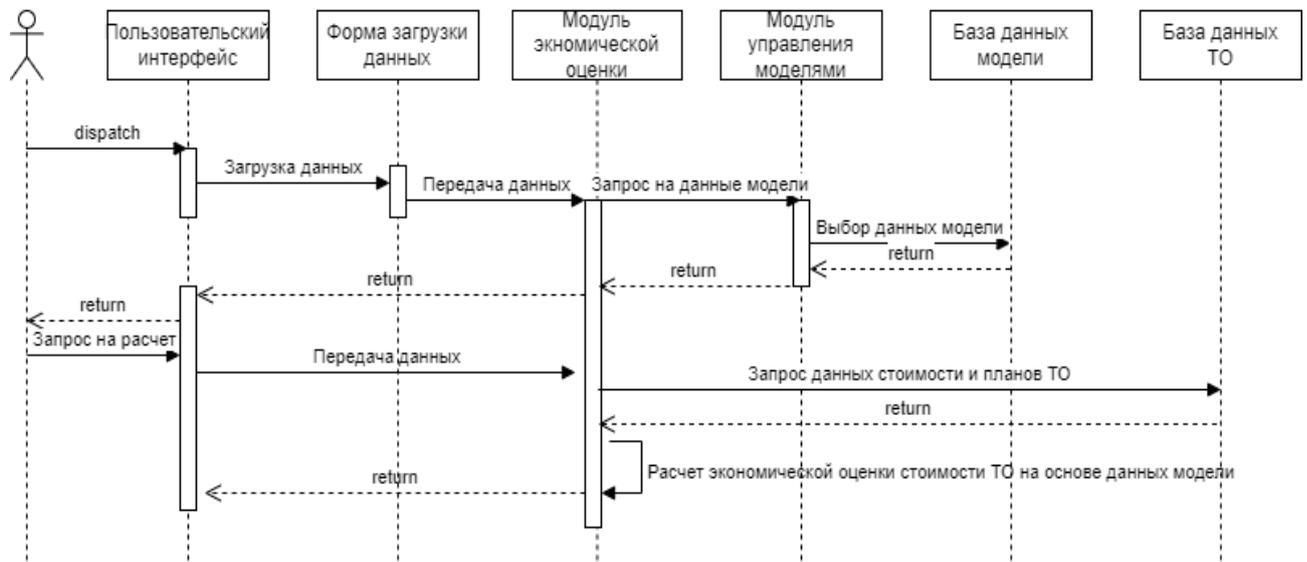


Рисунок 42 – Процесс получения экономических метрик (расходов/экономия) для ТОиР на основе анализа данных

Если для выбранного массива данных модель предсказывает высокую вероятность отказа оборудования (который может произойти до планового технического обслуживания и создать дополнительные расходы вследствие незапланированного простоя) или экономические потери при отказе, будет целесообразно провести внеплановое техническое обслуживание с предварительной корректировкой стратегии по обслуживанию и ремонту конкретной единицы оборудования. Основной метрикой для оценки выступает коэффициент рентабельности, позволяющий оценить, стоят ли дополнительные вложения в техническое обслуживание расчетной выгоды, которая будет получена посредством выполнения запланированного объема производственных задач.

Программный модуль принятия решений является агрегирующей надстройкой над подсистемами управления и анализом данных. С использованием алгоритмов анализа данных модуль позволяет прогнозировать возможные отказы оборудования и определять оптимальное время для проведения технического обслуживания. На основе полученных прогнозов, данных о расходах и последствиях отказов оборудования модуль представляет в рекомендательном формате перечень альтернативных вариантов, отвечающих на вопросы: когда и как проводить техническое обслуживание для оптимизации производственных процессов и снижения издержек.

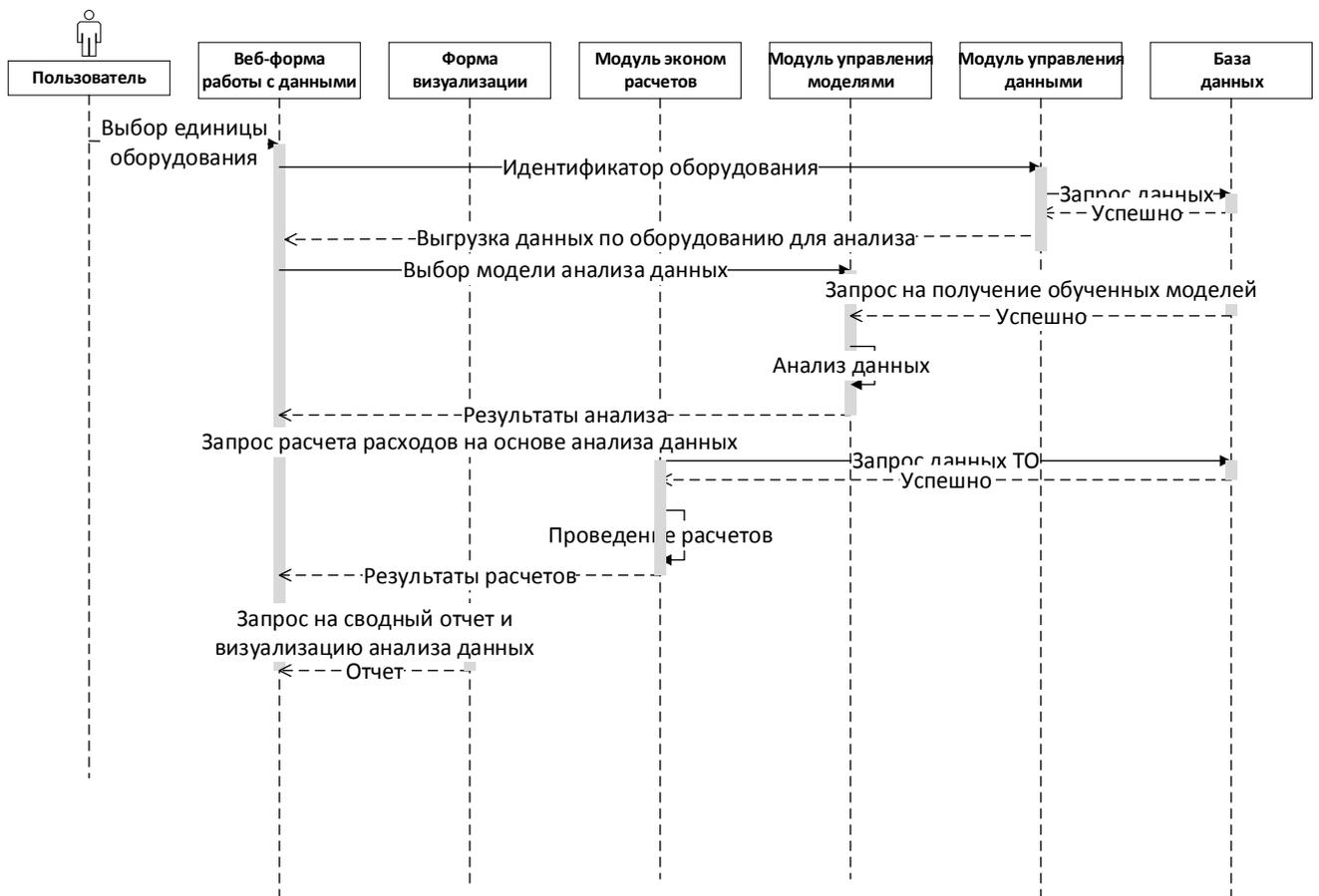


Рисунок 43 – Процесс принятия решений на основе данных

С использованием обратной связи от пользователей программной системы, модуль принятия решений позволяет отслеживать эффективность своих прогнозов и решений, для предоставления необходимых корректировок обновления модели

на основе новых данных и учтенного в базах данных опыта (истории технического обслуживания и ремонтных работ).

3.5. Выводы по ГЛАВЕ 3

На базе проблемно-ориентированного конвейера обработки данных промышленных роботов, реализовано программное обеспечение автоматизации процессов принятия решений по ТОиР в рамках предиктивного подхода в составе:

1) Модуля управления данными, предназначенного для загрузки данных из реляционного хранилища, формирования датасетов (в том числе, обучающих и тестовых выборок), оценку сходства датасетов методами Евклидова расстояния и динамической трансформации временной шкалы.

2) Модуля выявления аномальных значений, позволяющего сформировать диапазон допустимых значений при работе, что обеспечивает гибкость процесса мониторинга состояния промышленного робота по сравнению со штатными функциональными возможностями программного обеспечения мониторинга. Получено свидетельство о регистрации программы для ЭВМ «Диагностика аномальных состояний приводов промышленных манипуляторов» (Приложение 2).

3) Модуля управления моделями анализа данных, обеспечивающего процесс выбора, конструирования, обучения и тестирования модели для последующего использования и сохранения в нереляционное хранилище. Получено свидетельство о регистрации программы для ЭВМ «Модуль адаптивного выбора прогнозной модели работоспособности промышленного манипулятора» (Приложение 2).

4) Модуля поддержки принятия решений на основе данных, позволяющий оценить экономическую эффективность от внедрения предиктивной аналитики для ТОиР с использованием методов интеллектуального анализа данных.

4. Апробация программного обеспечения автоматизации процессов ТОиР с применением методов интеллектуального анализа данных

4.1. Характеристики наборов данных, использованных для апробации

Для апробации разработанной модели были взяты несколько массивов данных, сформированных по структуре БД, представленной на рисунке 26. Экспериментальная апробация модулей программного обеспечения была выполнена на основе задачи прогнозирования температуры привода промышленного робота. Точное и своевременное прогнозирование температуры электропривода позволяет: предотвращать его перегрев, оптимизировать работу и повысить энергоэффективность, определение оптимального момента для обслуживания или замены. Данные были получены в результате сбора при процессе выполнения однотипных операций на протяжении длительного промежутка времени (более 7 часов непрерывной работы).

Анализируемый набор данных состоит из четырех наборов данных (далее – датасетов). Каждый датасет содержит исторические данные о работе промышленного робота, каждый из которых суммарно выполнял однотипную рабочую операцию (перемещение паллеты, перемещение пластиковых кубов, работа со сварочным агрегатом, фрезерование поверхности) на протяжении 40 часов (таблица 1, 2). Для разбиения данных на обучающую и тестовую выборку выбраны пропорции 70:30 с применением рандомизации. График температуры каждого из электроприводов робота колеблется в определенном диапазоне (от 290 до 340 градусов по Кельвину).

Таблица 1 – Описание набора данных

Наименование пром. робота	Выполняемая задача	Номер эксперимента	Продолжительность эксперимента
Робот 1: Agilus («FANUC LRM- 200iD»)	Перемещение пластиковых кубиков внутри роботизированной ячейки	1	04:32:49
		2	12:56:36
		3	12:43:52
		4	09:52:36
Робот2: Frezer («KUKA KR 120 R2900»)	Фрезерование вазы из дерева	1	09:04:30
		2	12:16:04
		3	12:38:50
		4	09:37:23
Робот 3: Palitir («KUKA KR 40- PA»)	Перемещение на конвейер с паллеты и обратно коробок размерами 530×360×220 мм	1	07:40:20
		2	12:19:42
		3	12:19:42
		4	09:25:58
Робот 4: Svarka («ABB IRB1200»)	Сварочные работы	1	08:43:07
		2	12:18:59
		3	12:40:57
		4	09:45:26

Данные по каждому эксперименту содержат одинаковый набор признаков, собранных с датчиков подвижных частей промышленных роботов в ходе проведения серии экспериментов. Дискретность записи набора данных мониторинга составляет 1 секунду. Набор данных содержит 7 столбцов и ориентировочно 260-280 тысяч строк. Пример состава файла с данными представлен в приложении 7.

Таблица 2 – Описание признаков выборки

Обозначение	Наименование признака	Единица измерения	Описание
DATE_TIME	Время	-	Время замера показателей датчиков в формате HH:mm:ss
NUM_MOTOR	Номер привода	-	Номер подвижной части промышленного робота
CURR_ACT	Сила тока	мА	Сила тока
MOT_TEMP	Температура привода	С	Температура привода
TORQUE_AXIS_ACT	Момент силы	Н*м	Момент силы вращающейся части
VEL_AXIS_ACT	Угловая скорость	м/с	Угловая скорость вращения привода
VEL_AXIS	Линейная скорость	м/с	Линейная скорость вращения привода

4.2. Подготовка данных для апробации

На рисунке изображена матрица корреляции Пирсона для отобранных признаков (область значений [-1; 1]). Предварительно проведена нормализация данных (приложение 9). Далее следует оценка модуля значений корреляции Пирсона, так как признаков в исследуемом датасете относительно мало, то приемлемый порог является 0,95 (рисунок 44).

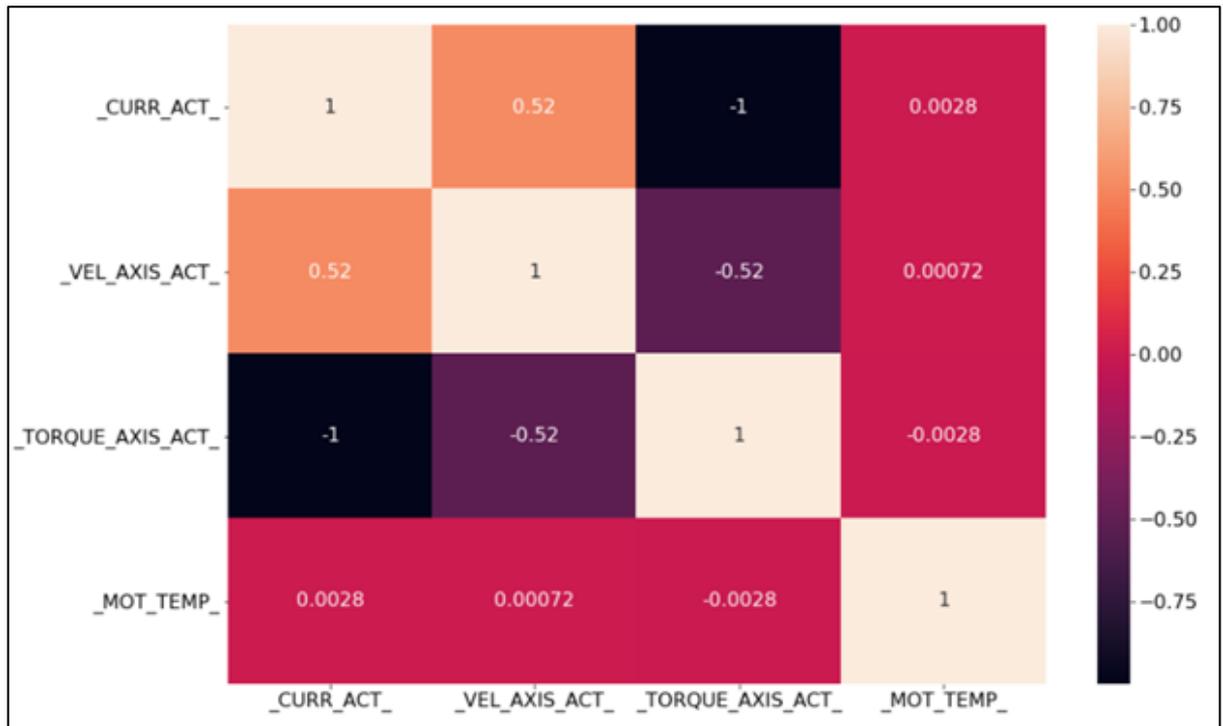


Рисунок 44 – Диаграмма корреляции признаков

Из матрицы следует, что «_CURR_ACT_» (сила тока) и «_TORQUE_AXIS_ACT_» (момент силы) сильно коррелированные признаки, следовательно, использование обоих в обучаемой модели не добавляет значимой информации для результата. Производится исключение признака «_TORQUE_AXIS_ACT_». «_CURR_ACT_» (сила тока) и «_VEL_AXIS_ACT_» (измеренная скорость вращения) имеют значение корреляции Пирсона 0,52, но так как исследуемых признаков мало, данные параметры можно оставить. Значение температуры слабо коррелирует с остальными признаками, для построения модели обязательно нужно использовать этот признак. После применения методов фильтрации было сформировано следующее множество признаков:

- сила тока;
- измеренная скорость вращения;
- температура привода.

Далее использован прямой метод последовательного выбора признаков, так как на предыдущем этапе был выделен основной, мало коррелирующий с остальными признак – температура привода. Данный признак обозначен как

целевой. Для сравнения результатов метода последовательного отбора признаков были проведены 3 серии опытов со следующими комбинациями признаков (таблица 3).

Таблица 3 – Комбинации признаков при отборе

№ опыта	Выбранные параметры
1	температура привода
	измеренная скорость вращения
2	температура привода
	сила тока
3	сила тока
	измеренная скорость вращения
	температура привода

Все опыты проводились с использованием набора данных о работе манипулятора KUKA KR 40-PA. В таблице 4 представлены результаты тестирования моделей. Критериями для сравнения являются: время обучения и точность прогнозирования целевого параметра модели.

Таблица 4 – Результаты тестирования моделей на разных наборах признаков

Модель обучения	№ опыта	T, сек	Кол-во эпох	Точность, %
Линейная модель	1	22,15	11	71,21
	2	6,61	4	80,65
	3	25,47	15	71,64
Многослойный перцептрон	1	5,79	3	71,26
	2	14,76	8	81,15
	3	9,55	5	71,84
Сверточная ИНС	1	21,23	12	68,73
	2	26,95	15	72,60
	3	29,81	16	70,74

Наибольшая точность предсказания второй серии экспериментов с комбинацией признаков под номером 2 (таблица 3). Найденное с использованием методов отбора подмножество признаков является оптимальным в рамках решения поставленной задачи. В приложении 8 представлена таблица со всеми комбинациями признаков и результатами тестирования моделей по каждой комбинации.

4.3. Результаты применения алгоритма оценки сходства сформированных выборок

Согласно выбранным в пункте 2.4 методам оценки сходства данных в таблицах 5, 6 представлены коэффициенты подобия данных между датасетами 4-х промышленных роботов.

Таблица 5 – Коэффициенты подобия данных (функция DTW)

Выборка	Робот 1	Робот 2	Робот 3	Робот 4
Робот 1	0,00	49,73	155,29	45,09
Робот 2	49,73	0,00	149,52	50,26
Робот 3	155,29	149,52	0,00	157,27
Робот 4	45,09	50,26	157,27	0,00

Таблица 6 – Коэффициенты подобия данных (функция расстояния Евклида)

Выборка	Робот 1	Робот 2	Робот 3	Робот 4
Робот 1	0,00	15,94	20,66	15,40
Робот 2	15,94	0,00	20,15	15,61
Робот 3	20,66	20,15	0,00	20,63
Робот 4	15,40	15,61	20,63	0,00

4.4. Обучение и тестирование базовых моделей

Методы анализа данных для тестирования функциональности разработанного программного обеспечения представлены в таблице 7.

Таблица 7 – Используемые модели для тестирования алгоритма

№	Модель	Описание
1	Linear	Линейная модель (с одним выходным слоем)
2	MLP	Многослойный перцептрон (два скрытых слоя по 64 нейрона и один выходной)
3	RNN	Рекуррентная нейронная сеть (размерность выходного пространства 32)
4	RFR	Регрессия случайного леса (количество деревьев 10, максимальная глубина 5)
5	GPR	Регрессор Гауссовского процесса

Результаты тестирования представлены в таблице 8.

Таблица 8 – Результаты тестирования моделей анализа данных

	Линейная модель (Linear)		Многослойный перцептрон (MLP)	
	Время, сек	Точность, %	Время, сек	Точность, %
Робот 1	8,67	85,68	3,84	83,05
Робот 2	16,43	79,65	4,3	80,75
Робот 3	14,25	86,90	3,15	76,04
Робот 4	14,57	78,95	4,63	71,74
	Рекуррентная нейронная сеть (RNN)		Регрессор случайного леса (RFR)	
	Время, сек	Точность, %	Время, сек	Точность, %
Робот 1	10,91	82,80	7	98,1
Робот 2	11,33	82,74	5	89,87
Робот 3	11,51	80,21	5	93,43
Робот 4	9,14	75,72	9	90,29
	Регрессор Гауссовского процесса (GPR)			
	Время, сек	Точность, %		
Робот 1	40	97,9		
Робот 2	215	91,31		
Робот 3	152	95,51		
Робот 4	188	93,11		

4.5. Тестирование моделей на выборках, сформированных с парка роботов

В таблице 9 представлены результаты тестирования разных моделей машинного обучения и ИНС на выборках с роботов. Критерии оценки – время работы модели и точность прогнозируемых параметров.

Таблица 9 – Результаты тестирования моделей на выборках с разных роботов

Тестовая Обучающая	Робот 1		Робот 2		Робот 3		Робот 4	
	Время	Точность	Время	Точность	Время	Точность	Время	Точность
Линейная модель (Linear)								
Робот 1	8,67	70,83	2,01	47,79	2,17	51,87	2,03	59,37
Робот 2	2,01	44,79	16,43	51,04	2,57	44,79	2,12	43,75
Робот 3	1,99	51,87	2,00	30,20	14,25	78,12	2,29	26,04
Робот 4	2,31	28,12	2,61	43,75	2,12	37,5	14,57	42,71
Многослойный перцептрон (MLP)								
Робот 1	3,84	73,96	2	61,45	2	64,53	2	66,66
Робот 2	2	38,54	4,30	50,0	2	53,44	2	49,64
Робот 3	2	22,91	2	28,15	3,15	76,04	2	27,08
Робот 4	2	36,45	2	42,70	2	44,71	4,63	38,54
Рекуррентная нейронная сеть (RNN)								
Робот 1	10,91	67,71	2	58,33	2	59,37	2,12	63,54
Робот 2	2	35,41	11,33	60,42	2	41,66	2	34,37
Робот 3	2	35,41	2	31,25	11,51	80,21	2	31,25
Робот 4	2	41,66	2	39,58	2	36,45	9,14	42,71
Регрессор случайного леса (RFR)								
Робот 1	7	98,10	2,32	88,26	2,42	88,19	2,11	84,45
Робот 2	2,23	19	5	89,87	2,41	29,18	1,9	89,80
Робот 3	1,92	14,74	2,76	87,78	5	93,43	2,10	86,78
Робот 4	2,22	16,09	1,98	87,53	2,22	36,59	9	90,29
Регрессор Гауссовского процесса (GPR)								
Робот 1	40	97,9	12,11	87,53	22,99	84,87	32,58	82,2
Робот 2	12,84	40,64	215	91,31	12,16	62,5	13,4	71,46
Робот 3	11,91	45,22	12,72	55,69	152	95,51	22,45	76,07
Робот 4	12,01	31,43	12,13	41,53	22,15	35,54	188	93,11

Наклонным шрифтом (по главной диагонали) выделены результаты тестирования, при которых обучающая и тестовая выборка взяты с одного и того же промышленного робота. Жирным шрифтом выделены числовые значения, модели которых приемлемы для использования при анализе данных (основываясь на метриках); данные модели показывают достаточную точность как для тестовой

выборки, сформированной с оригинальной единицы оборудования (т.е. тестовая и обучающая выборка с одного набора данных) так и для тестовой выборке, сформированной из набора данных, полученного в процессе работы другой производственной единицы, при этом с наименьшими потерями точности, относительно других тестируемых пар. Наименьшее падение точности показывают модели 1, 2, 3. В случаях, когда этого не происходит, следующей по точности являются модели 4, 5, обучающая выборка которых следующая по сходству с данными Робот 1 – Робот 2 (рисунок 45).

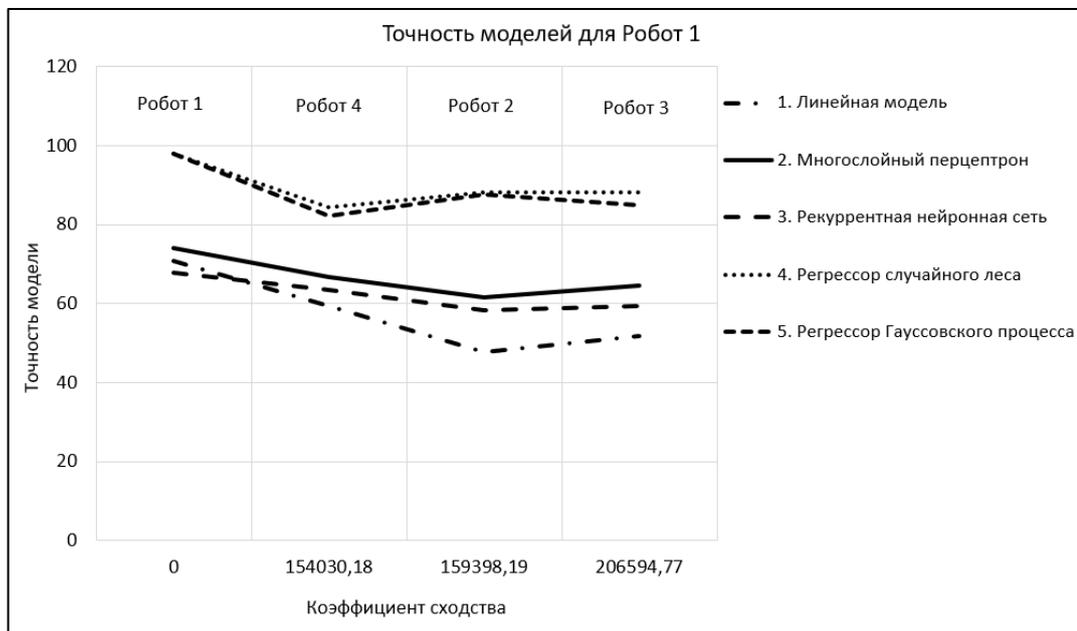


Рисунок 45 – Зависимость точности прогнозирования моделей от коэффициента сходства данных между Робот 1 (обучающая выборка) и Робот 2-4 (тестовые выборки)

Описанная тенденция сохраняется на все типах промышленных роботов, кроме Робот 4, где меньшее падение точности прогнозирования показывали модели (1, 3–5) на второй тестовой выборке по коэффициенту сходства (рисунок 46). Результаты прогнозирования целевого параметра (температуры привода) и зависимость точности модели от коэффициента сходства данных представлены в приложении 10.

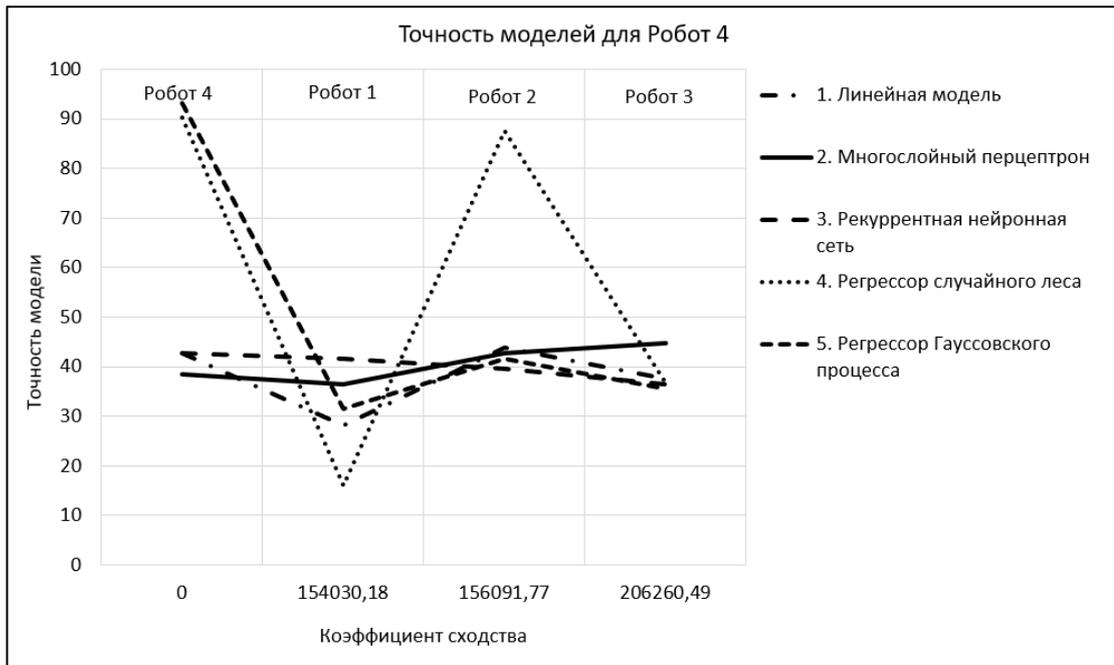


Рисунок 46 – Зависимость точности прогнозирования моделей от коэффициента сходства данных между Робот 4 (обучающая выборка) и Робот 1-3 (тестовые выборки)

С использованием выбранного в пункте 2.3.4. метода улучшения метрик моделей машинного обучения (бустинг) получены следующие значения (таблица 10). Модели, использующие алгоритм регрессии (случайный лес и Гауссовский процесс) показали достаточно хорошие результаты, следовательно, улучшение этих моделей не требуется.

Таблица 10 – Результаты тестирования моделей с учетом этапа улучшения методом бустинга

Обучающая/тестовая	Робот 1	Робот 2	Робот 3	Робот 4
Линейная модель (Linear)				
Робот 1	85,68	81,12	72,40	68,14
Робот 2	74,45	79,63	70,61	70,47
Робот 3	75,46	81,65	86,93	74,06
Робот 4	77,91	70,95	74,77	79,01
Многослойный перцептрон (MLP)				
Робот 1	83,01	80,51	74,50	79,83
Робот 2	75,74	80,75	72,96	72,57
Робот 3	77,32	81,71	73,28	73,91
Робот 4	73,16	80,43	75,37	71,73
Рекуррентная нейронная сеть (RNN)				

Робот 1	83,95	81,23	74,27	76,32
Робот 2	79,21	82,78	77,15	70,81
Робот 3	71,83	79,51	80,14	72,84
Робот 4	80,45	72,11	79,32	75,68

После применения метода улучшения метрик модели процент «ухудшения» точности прогнозирования на тестовых выборках, сформированных на прочих единицах оборудования, снизился до 9–12% в среднем. На рисунке 44 представлена зависимость точности прогнозирования модели от коэффициента сходства данных (между обучающей выборкой с одной единицы оборудования и тестовыми с прочих единиц).

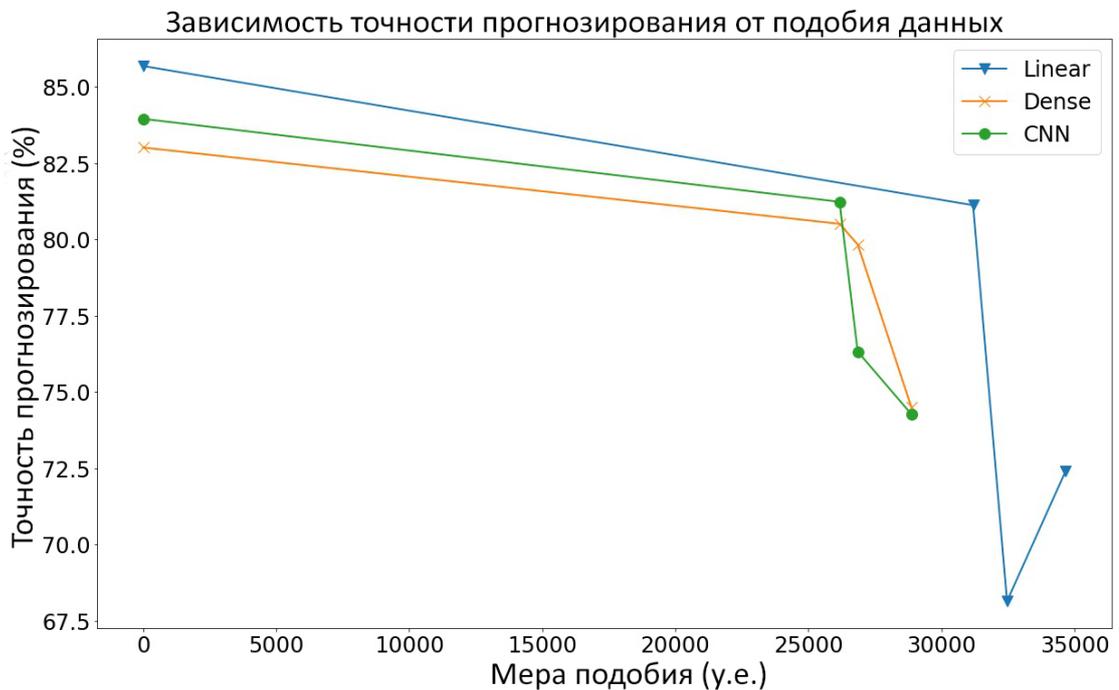


Рисунок 47 – Зависимость метрик тестирования от меры сходства данных (методом DTW)

На рисунке 45 представлены результаты прогнозирования температуры привода одной из моделей (обученной на данных с робота «agilus» и протестированной на данных с робота «svarka»).

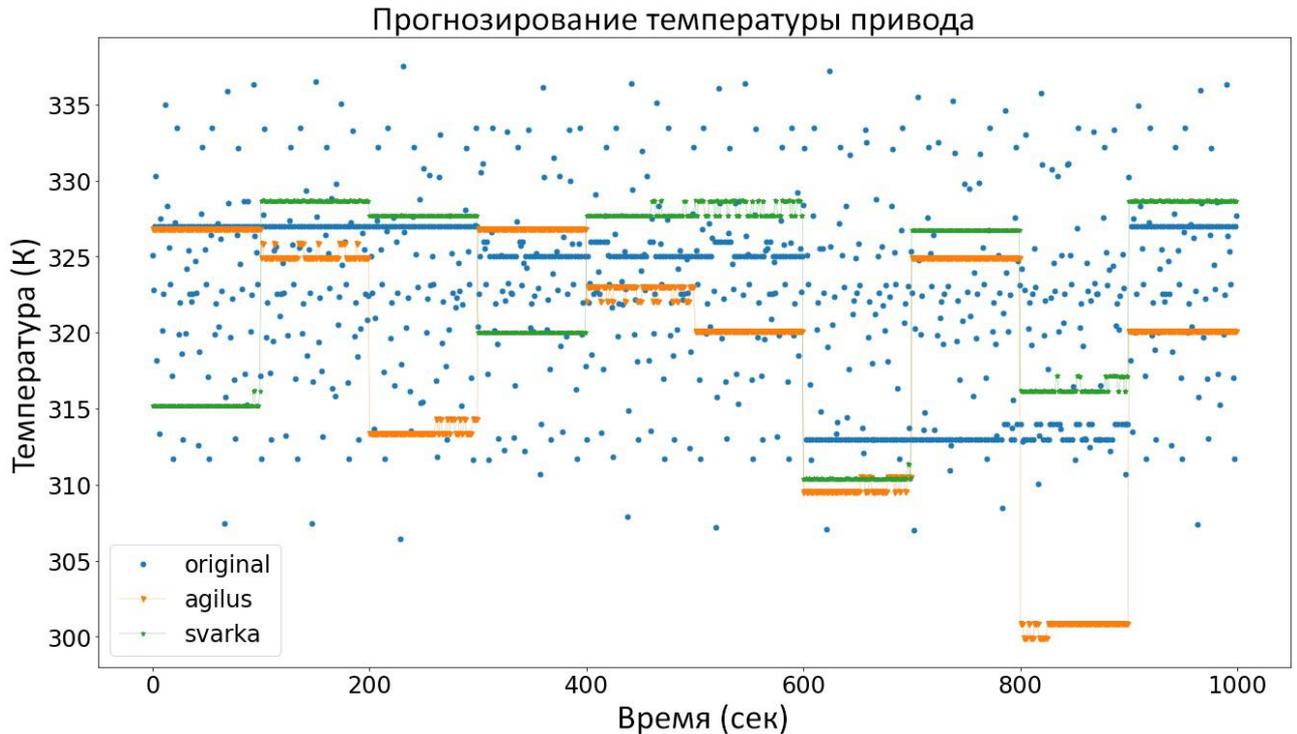


Рисунок 48 – Пример прогнозирования температуры привода моделью, обученной на работе «agilus»

Точность прогнозирования целевого параметра (температуры привода) промышленного робота с использованием моделей подобранных алгоритмом адаптивного выбора моделей машинного обучения ниже чем при использовании модели обученном на исходном наборе данных и составляет допустимые 30% (при доверительном интервале $\pm 0,5$ градусов прогнозирования температуры привода – рисунок 40), при этом время предсказания сокращается в среднем на 60-70 %, а потери точности от 5 до 20%. Для моделей регрессор случайного леса, рекуррентной нейронной сети, регрессора Гауссовского процесса и многослойного перцептрона применение разработанного алгоритма для прогнозирования значений с допустимой погрешностью является целесообразным для выигрыша во времени решении задачи при небольшой потере точности. Результаты тестирования полученных моделей могут быть использованы в качестве обратной связи для: настройки параметров модели, критерием валидации использования модели для решения поставленной задачи.

Структурно-функциональная схема управления состоянием промышленного манипулятора на основе результатов прогнозирования температуры электропривода представлена на рисунке 49.



Рисунок 49 – Структурно-функциональная схема управления состоянием промышленного манипулятора на основе результатов прогнозирования температуры электропривода

Выходные данные из системы мониторинга являются объектом изучения для прогнозирования остаточного срока службы электропривода на оси манипулятора и прогнозирования потенциальных отказов в рамках процесса предиктивного технического обслуживания и ремонта. Методы статистического и интеллектуального анализа являются неотъемлемой частью предиктивного анализа. Данные, полученные в результате анализа и имеющие удовлетворительную точность для поставленной задачи, используются для процесса принятия решений при техническом обслуживании и ремонта с экономической оценкой возможных альтернатив (инспекция, настройка, ремонт

или замена оборудования) в зависимости от: стоимости операции, наличия комплектующих, целесообразности проведения тех или иных работ. На основе принятого решения формируется план, включающий в себя предписания по техническому обслуживанию и ремонту. Руководствуясь предписаниями по техническому обслуживанию, исполнительными устройствами формируются управляющие воздействия U , представляющие из себя плановые или корректирующие операции. Точное прогнозирование температуры позволяет своевременно выявить риски перегрева электропривода, что критически важно для предотвращения сбоев в работе робота. Это позволяет минимизировать простои и повысить надежность производственного процесса. Знание прогнозируемой температуры позволяет оптимизировать график технического обслуживания электропривода, что продлевает его срок службы. Также прогноз температуры позволяет настроить режимы работы робота для повышения эффективности производительности, потребления энергии и минимизации тепловых нагрузок.

4.6. Оценки эффективности применения разработанного программного обеспечения

Программный модуль оценки экономической целесообразности принятия решений для предиктивного технического обслуживания основан на нескольких ключевых параметрах:

- Расходы на правильно спрогнозированный отказ – стоимость, связанная с необходимостью замены или починки оборудования в случае, если предиктивные модели верно определили возможный отказ.
- Расходы на обслуживание, включающие расходы на регулярное техническое обслуживание оборудования, в том числе: проверки, замены компонентов и другие операции.

- Расходы на обучение моделей анализа – затраты на обучение и настройку моделей предиктивного анализа для определения вероятности отказа оборудования.

- Общая экономия – оценка суммарной выгоды от применения предиктивного технического обслуживания, включая сокращение потерь от непредвиденных отказов, увеличение эффективности обслуживания и улучшение производственных процессов.

Для парка промышленных роботов разработана следующая экономическая модель принятия решений (на основе исследованных подходов из пункта 2.5):

$$TP_{savings} = \sum_{i=\#TP} (R - r) \left(\frac{\Delta t_i}{n} \right) \quad (19)$$

$$FN_{cost} = \#FN \times R^F FP_{cost} \quad (20)$$

$$Total_{savings} = \sum_{robot} \sum_{node} (TP_{savings} - FN_{cost} - FP_{cost}), \text{ где} \quad (21)$$

n – временной промежуток для прогнозирования,

Δt_i – временной промежуток для анализа выборки,

R – стоимость замены комплектующих,

r – стоимость ремонта оборудования.

TP – истинные прогнозы отказов оборудования в рамках временного промежутка.

FN – прогнозы отказов оборудования, не обнаруженные в рамках временного промежутка.

$TP_{savings}$ – оценка экономии на основе обнаруженных сбоев,

FN_{cost} – оценка расходов стоимости замены оборудования на основе не выявленных сбоев,

FP_{cost} – оценка расходов стоимости проверки оборудования на основе ложных данных анализа,

$Total_{savings}$ – итоговая оценка расходов стоимости технического обслуживания на основе анализа данных.

С учетом расходов не только на оценку прогнозных моделей, но и на ресурсы, необходимые для конструирования и обучения получается следующая экономическая модель:

$$TP_{savings} = \sum_{i=\#TP} (R - r) \left(\frac{\Delta t_i}{n} \right) - (t_k * Cost_k + t_{test} * Cost_{test}) \quad (22)$$

t_k – временные затраты на конструирование модели интеллектуального анализа данных.

$Cost_k$ – оценка расходов, необходимых для процесса конструирования модели интеллектуального анализа данных.

t_{test} – временные затраты на тестирование и валидацию модели интеллектуального анализа данных.

$Cost_{test}$ – оценка расходов, необходимых для процесса тестирования и валидацию модели интеллектуального анализа данных.

Расходы на конструирование и тестирование основаны на следующих факторах: опыт специалистов, объем и качество данных, используемые инструменты и инфраструктура, а также сложность задачи. Как правило, затраты включают в себя трудозатраты специалистов, стоимость оборудования и программного обеспечения, а также ресурсы для обучения и развертывания модели. В общем виде эти затраты включают в себя следующие операции:

- Подготовка данных: сбор, очистка, преобразование и анализ данных.
- Выбор модели: определение наилучшего типа модели для конкретной задачи.
- Обучение модели: использование алгоритмов машинного обучения для настройки модели на обучающих данных.
- Тестирование модели: оценка качества модели с использованием метрик и тестирование на отложенной выборке.
- Тюнинг модели: настройка гиперпараметров для улучшения производительности модели.

- Развертывание модели: интеграция модели в рабочее окружение и тестирование на реальных данных.

С учетом разработанного комплекса алгоритмов по подбору моделей анализа на основе оценки сходства данных, интегрированного в конвейерную информационную инфраструктуру управления моделями интеллектуального анализа, получается следующая экономическая модель:

$$TP_{savings} = \sum_{i=\#TP} (R - r) \left(\frac{\Delta t_i}{n} \right) - (t_{test} * Cost_{test}) \quad (23)$$

$$FN_{cost} = FN \times R^F FP_{cost} * k_{accuracy} + (t_{select} * Cost_{select}) \quad (24)$$

t_{test} – временные затраты на тестирование и валидацию модели интеллектуального анализа данных.

$Cost_{test}$ – оценка расходов, необходимых для процесса тестирования и валидацию модели интеллектуального анализа данных.

$k_{accuracy}$ – коэффициент потери точности модели интеллектуального анализа данных.

t_{select} – временные затраты на подбор и рекомендацию модели интеллектуального анализа данных.

$Cost_{select}$ - оценка расходов, необходимых для процесса подбора и рекомендации модели интеллектуального анализа данных.

Помимо расходов (временных, вычислительных и трудозатрат) на подбор подходящей модели интеллектуального анализа новых данных, отдельным аспектом выделяется переменная $k_{accuracy}$, с помощью которой производится поправка на возможное ухудшение результатов тестирования, что в свою очередь повышает FN_{cost} на основе прогнозирования целевых данных.

В случае выбора опции поиска подходящей модели анализа под имеющиеся данные, можно пренебречь расходами на конструирование новых моделей. Использование уже обученных моделей машинного обучения на новых данных с высокой степенью сходства с обучающей выборкой позволяет экономить время и

ресурсы на повторном обучении модели. Данный подход снижает затраты на вычислительные мощности и трудозатраты специалистов, необходимые для проведения процесса конструирования заново. Также использование уже обученной модели позволяет сократить временные затраты этапы развертывания и внедрения модели в производственную среду.

4.7. Разработка имитационных моделей для расчета экономической эффективности

Для оценки экономии временных ресурсов на обучение моделей интеллектуального анализа, посредством внедрения предложенного алгоритма подбора моделей на основе оценки сходства данных сформирована имитационная модель (рисунок 50). В программной среде моделирования Arena Simulation разработана модель для сравнения затраченного времени. Производится сравнение двух величин: «тестирование модели» и «обучение + тестирование модели». Данное сравнение этих величин наглядно демонстрирует разницу во временных затратах.

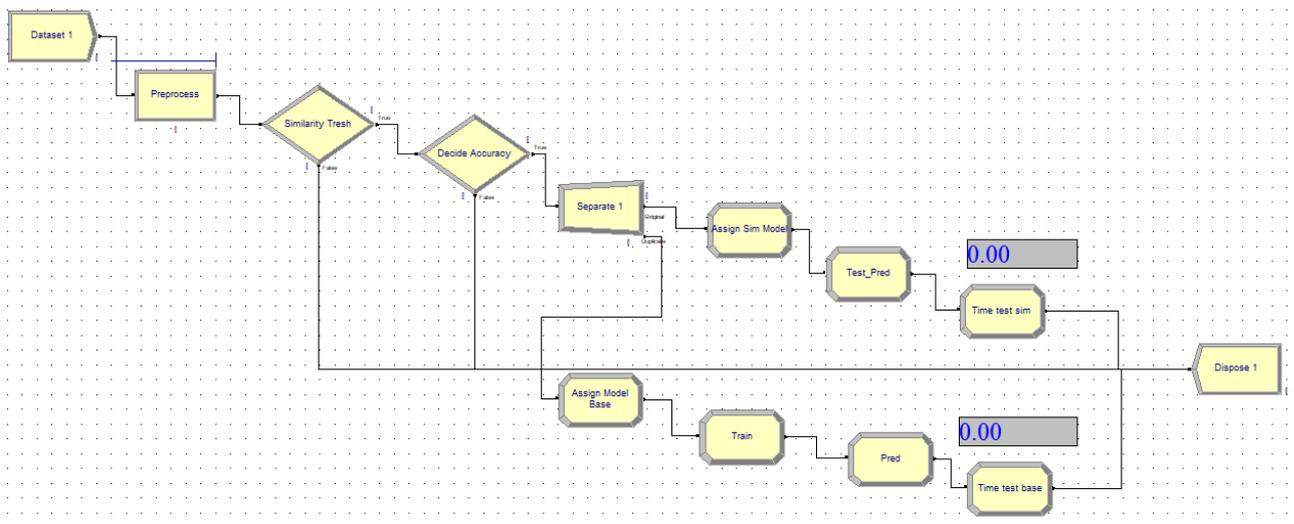


Рисунок 50 – Имитационная модель экономии временных ресурсов при использовании подхода подбора обученных моделей

Имитационная модель состоит из следующих блоков:

- Блок «Создание сущности» – под сущностью подразумевается набор данных (датасет), который выбран для анализа. Для примера определено создание 1 сущности в 10 минут при 24 часовом моделировании.
- Блок «Подготовка данных» – в данном блоке предполагается моделирование процесса оценки сходства данных, который занимает от 5 до 15 секунд (при условии равномерного распределения).
- Условный оператор 1 – данный блок является «фильтром» для отсеивания датасетов, не удовлетворяющих условиям для дальнейшей работы (то есть значение сходства данных ниже порогового). В процессе имитационного моделирования использовались значения от 5 до 90% (чем выше процент – тем выше найти вероятность подходящего для анализа датасета).
- Условный оператор 2 – данный блок является «фильтром» для отсеивания условных моделей анализа, показавших неудовлетворительную точность (то есть ниже порогового значения). В процессе имитационного моделирования использовались значения от 5 до 50% (чем выше процент – тем менее точная модель проходит отбор).
- Блок «Разделение сущности» – данный блок производит распараллеливание процесса для сравнения двух переменных – тестирование модели анализа данных и тестирование новой модели анализа данных с учетом времени её обучения. Время обучения модели варьируется от 40 до 215 секунд (при условии равномерного распределения), время тестирования модели варьируется от 12 до 32 секунд (при условии равномерного распределения).
- Блок «Уничтожение сущности» – данный блок является стандартным для большинства моделей в среде Arena. Блок выполняет функцию уничтожения сущностей, а также подсчет этих сущностей для формирования итогового сводного отчета по результатам моделирования.

Результаты имитационного моделирования представлены в таблице 11. Данные для моделирования взяты из таблиц 8, 9, 10.

Таблица 11 – Результаты имитационного моделирования экономии временных ресурсов

Процент датасетов	Процент пороговой потери точности модели									
	5	10	15	20	25	30	35	40	45	50
5	0,0%	0,0%	0,0%	5,6%	5,6%	8,5%	8,5%	8,5%	8,5%	8,5%
10	0,0%	0,0%	12,6%	9,7%	9,8%	12,6%	12,9%	12,9%	12,9%	12,9%
15	8,3%	7,3%	7,5%	7,5%	9,2%	11,7%	11,7%	11,7%	12,2%	12,2%
20	8,3%	7,3%	7,8%	7,5%	8,2%	8,4%	8,4%	8,4%	8,9%	9,7%
25	4,8%	5,2%	6,9%	7,6%	8,4%	9,1%	9,1%	9,1%	9,4%	9,9%
30	8,5%	8,6%	10,3%	9,5%	9,8%	9,9%	9,3%	9,3%	9,6%	10,2%
35	11,3%	10,1%	10,8%	10,3%	10,5%	9,8%	9,7%	9,7%	9,7%	10,3%
40	10,5%	8,3%	10,9%	10,5%	10,5%	11,3%	11,3%	11,3%	11,2%	11,6%
45	10,5%	9,1%	11,4%	11,9%	11,6%	11,6%	11,6%	11,7%	11,6%	12,0%
50	12,2%	10,5%	12,9%	12,3%	12,6%	11,7%	11,7%	11,8%	11,7%	11,9%
55	11,2%	8,9%	13,0%	12,4%	11,8%	12,2%	12,4%	11,9%	12,1%	12,3%
60	11,5%	10,9%	12,3%	11,5%	12,1%	13,1%	12,4%	12,0%	12,2%	12,2%
65	11,5%	12,0%	13,7%	12,4%	11,8%	12,2%	11,4%	11,5%	11,4%	11,5%
70	16,1%	15,8%	16,4%	13,7%	12,8%	12,4%	12,3%	12,3%	12,3%	12,5%
75	18,1%	16,4%	15,7%	12,9%	12,2%	11,9%	11,9%	12,3%	12,0%	12,7%
80	13,1%	14,2%	14,6%	12,7%	11,9%	11,9%	12,1%	12,1%	11,8%	12,4%
85	18,6%	16,1%	15,0%	12,0%	12,0%	11,9%	12,2%	12,2%	12,0%	12,4%
90	14,2%	10,8%	12,9%	11,7%	11,9%	11,9%	12,2%	11,9%	11,9%	12,4%

Процентные данные для таблицы получены из следующих расчетов: (время тестирования модели) / (время обучения модели + время тестирования модели) * 100%.

Визуализация данных представлена на рисунке 51. Исходя из полученных данных, можно сделать следующий вывод: максимальная экономия временных ресурсов возможна, если имеется большой объем схожих данных, из которых на сформированных датасетах производится наиболее точное тестирование моделей анализа; минимальная экономия (или ее отсутствие) можно увидеть при условии малого наличия схожих данных и низкого порога точности моделей анализа.

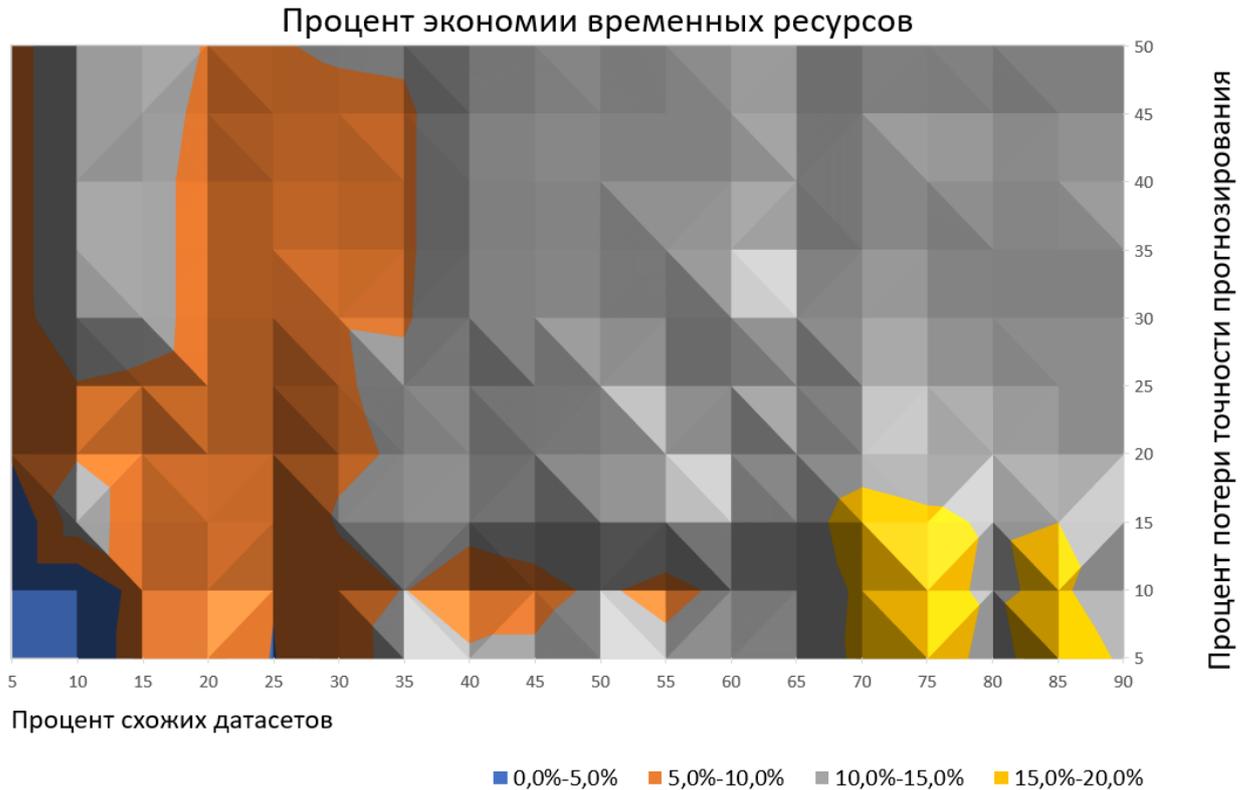


Рисунок 51 – Визуализация результатов имитационного моделирования экономии временных ресурсов

Для оценки экономических затрат для принятия решений на основе данных сформирована следующая имитационная модель (рисунок 52). Данная модель учитывает расходы не только на оценку прогнозных моделей, но и на ресурсы, необходимые для их конструирования и обучения. Имитационная модель состоит из следующего набора блоков:

- Блок «создание сущности» – под сущностью подразумевается модель анализа данных, на основе прогнозов которой производятся экономические расчеты для принятия дальнейших решений по инспекции, ремонту и обслуживанию оборудования. Определено создание 1 сущности раз в 10 минут при 24 часовом моделировании.
- Блок для хранения переменных, в котором производится первичная инициализация расчетных значений, которые включают в себя: время конструирования и тестирования модели (40-215 и 12-32 секунд соответственно), условные стоимости конструирования и тестирования (9-11 у.е. и 7-10 у.е.

соответственно), условная экономия за счет корректно спрогнозированных отказов (10000-20000 у.е.).

- Блоки подготовки данных, конструирования и тестирования модели, использующие значения из ранее инициализированных переменных.
- Блоки расчетов расходов на основе данных, учитывающий время и стоимость работы с моделями анализа данных, расходы по инспектированию оборудования (4000 у.е.) и ремонту оборудования (8000 у.е.).
- Блок итогового расчета полученной условной выгоды от принятия решений на основе данных.
- Блок калькуляции среднего значения итогового расчета в течении всего времени работы имитационной модели.

Результатом работы модели является 11622 у.е. итоговая средняя экономия (моделирование 24 часа, обработано 144 сущности).

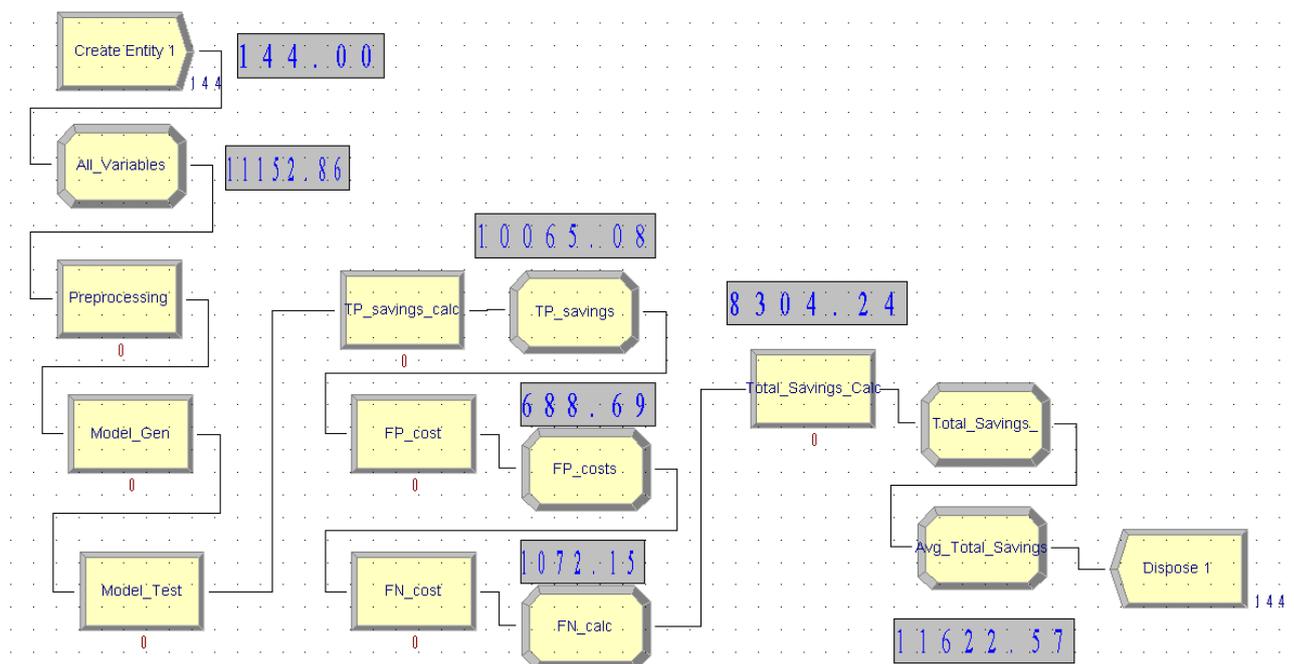


Рисунок 52 – Имитационная модель (формулы 20-22)

Для моделирования процесса оценки экономических затрат с учетом разработанного алгоритма подбора моделей интеллектуального анализа сформирована следующая модель (рисунок 53).

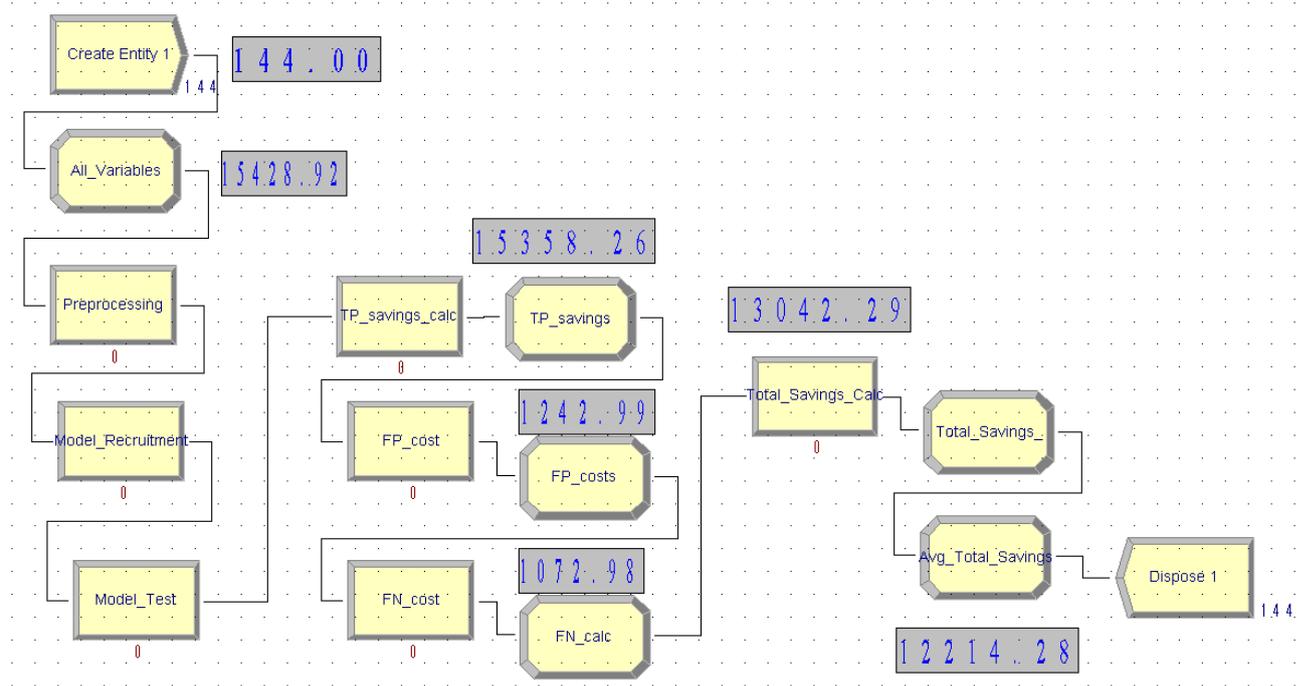


Рисунок 53 – Имитационная модель (формулы 21, 23, 24)

В данной имитационной модели заменен блок конструирования модели анализа данных на блок подбора модели. Добавлены переменные для времени и стоимости процесса подбора модели (20–30 секунд и 5–10 у.е. соответственно). Также были убраны переменные, связанные с конструированием моделей. В блок FP_costs добавлены дополнительные расходы на подбор модели и коэффициент потенциальной потери точности прогнозирования (в пределах 8–25% при условии равномерного распределения). Результатом работы модели является 12214 у.е. итоговая средняя экономия (моделирование 24 часа, обработано 144 сущности).

Исходя из результатов сравнения имитационных моделей, при заданных параметрах моделирования, вторая модель (спроектированная с учетом разработанной алгоритмической модели по подбору моделей интеллектуального анализа на основе оценки сходства данных) позволяет сократить расходы при принятии решений на основе данных для ТОиР на 5,10 %.

4.8. Внедрение результатов диссертационного исследования

Результаты работы используются в ООО «Инспайр-технологии» (г. Томск) – в составе программного обеспечения мониторинга работоспособности и оценки текущего состояния промышленного манипулятора. Акт о внедрении приложен к диссертационной работе [Приложение 12]. По результатам разработки отдельных модулей информационной системы получено свидетельство о регистрации программы для ЭВМ «Диагностика аномальных состояний приводов промышленных манипуляторов» [Приложение 1].

Функционал разработанного информационного обеспечения основан на анализе физических параметров сервоприводов с учетом косвенной информации, получаемой с контроллера промышленного манипулятора. Штатное программное обеспечение не анализирует возможный выход из строя оборудования, оно лишь выводит значения измеряемых величин на монитор и подсвечивает критические значения величин, установленные заводом-изготовителем. Разработанный в рамках диссертационной работы комплекс алгоритмов позволяет сформировать диапазон допустимых значений при работе, обеспечивающий гибкость процесса мониторинга состояния оборудования и корректность соблюдения технологического процесса роботизированной ячейки по сравнению со штатным средством мониторинга. Алгоритм используется для отслеживания расхождений при перемещении грузов (в частности – паллет) с формированием оповещений на интерфейс оператора о возникающих отклонениях в рабочем цикле манипулятора [Приложение 11].

Внедренные в информационную инфраструктуру предприятия, программные модули объединены в единое решение по принципу SaaS (Software as a Service). Объединение модулей (включающих веб-интерфейс для специалиста по анализу данных) в прототип информационной системы обеспечивает набор функций по работе и управлению моделями анализа данных для парка промышленных роботов.

4.9. Выводы по ГЛАВЕ 4

Результаты апробации разработанного программного обеспечения поддержки принятия решений при планировании технического обслуживания и ремонта оборудования на базе проблемно-ориентированного конвейера обработки данных позволяют сделать следующие выводы:

1) Разработанный алгоритм вычисления меры сходства наборов исторических данных о работе промышленного робота обеспечивает подбор модели анализа данных с потерей точности не более 12%.

2) Разработанный алгоритм рекомендации и оценки эффективности модели машинного обучения для прогнозирования параметров работы промышленного робота, обеспечивает сокращение временных ресурсов на использовании модели анализа данных до 18%, при пороговом значении потери точности до 12%.

3) Разработанный алгоритм поддержки принятия решений по ТОиР позволяет повысить экономическую эффективность на 5,10% за счет экономии временных ресурсов на обучение новых моделей анализа данных.

Заключение

1) Предложена методика проектирования информационной системы для автоматизации предиктивного технического обслуживания и ремонта промышленных роботов на базе проблемно-ориентированного конвейера обработки данных, включающую комплекс алгоритмов по обработке данных разнородного оборудования и управления моделями анализа данных. Результаты имитационного моделирования продемонстрировали сокращение экономических расходов на информационное обеспечение процесса принятия решения при ТОиР на 5,1% за счет экономии временных ресурсов на конструирование новых моделей анализа данных.

2) Разработан алгоритм автоматизированного подбора и оценки эффективности модели машинного обучения для прогнозирования параметров работы промышленного робота, обеспечивающий сокращение временных ресурсов на использование моделей анализа данных до 18%.

3) Разработан алгоритм вычисления меры сходства наборов исторических данных о работе промышленного робота, обеспечивающий подбор модели анализа данных с потерей точности прогнозирования не более 12%.

4) Разработан алгоритм автоматизированного поиска аномальных значений параметров промышленных манипуляторов, обеспечивающий гибкость процесса мониторинга состояния за счет формирования диапазона допустимых значений параметров работы на каждой операции рабочей программы по сравнению со штатными функциональными возможностями программного обеспечения мониторинга.

5) Разработан прототип программного обеспечения, выполняющий функцию управления и подбора моделей анализа данных промышленных манипуляторов в информационной инфраструктуре ООО «Инспайр-технологии». Программное обеспечение осуществляет гибкость процесса мониторинга состояния оборудования и производственного технического процесса, а также

визуализацию результатов анализа данных работы электроприводов промышленных манипуляторов.

Список сокращений и условных обозначений

ИНС (Искусственная нейронная сеть) – математическая модель, а также ее программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей

СУБД (Система управления базами данных) – совокупность программных средств, обеспечивающих управление и создание баз данных

ТОиР Техническое обслуживание и ремонт

Framework Программное обеспечение, позволяющее автоматизировать разработку тестирование программного продукта

MongoDB Документоориентированная система управления базами данных с открытым исходным кодом, классифицирована как NoSQL

NoSQL Ряд подходов, направленных на реализацию СУБД

Python Высокоуровневый язык программирования общего назначения

AutoML Автоматическое машинное обучение

DT Цифровой двойник

LSTM Сети долгой краткосрочной памяти

MLP Многослойный перцептрон

RNN Рекуррентная нейронная сеть

RFR Регрессор случайного леса

GPR Регрессор Гауссовского процесса

CBR Подход на основе прецедентов

CPS Кибер-физическая система

MA Обслуживание на основе аналитики

Список литературы

1. Do P., Iung B., Cavalcante C. Reliability and maintenance cost forecasting for systems with multistate components using artificial neural networks //2019 4th International Conference on System Reliability and Safety (ICSRS). – IEEE, 2019. – С. 181-185.
2. Hanshuo M. et al. Research on fault prediction method of electronic equipment based on improved SVR algorithm //2020 Chinese Automation Congress (CAC). – IEEE, 2020. – С. 3092-3096.
3. Rathaur S., Kamath N., Ghanekar U. Software defect density prediction based on multiple linear regression //2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA). – IEEE, 2020. – С. 434-439.
4. Alhirmizy S., Qader B. Multivariate time series forecasting with LSTM for Madrid, Spain pollution //2019 international conference on computing and information science and technology and their applications (ICCISTA). – IEEE, 2019. – С. 1-5.
5. Trivedi S. et al. Predictive maintenance of air conditioning systems using supervised machine learning //2019 20th International Conference on Intelligent System Application to Power Systems (ISAP). – IEEE, 2019. – С. 1-6.
6. Sun H. H. et al. Multiaxial Fatigue Life Prediction Based on Critical Plane Method //2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE). – IEEE, 2019. – С. 888-900.
7. Solanki R., Ravilla S. H., Bein D. Study of distributed framework hadoop and overview of machine learning using apache mahout //2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC). – IEEE, 2019. – С. 0252-0257.
8. Özkan S. B. et al. Comparison of open source data mining tools: Naive Bayes algorithm example //2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT). – IEEE, 2019. – С. 1-4.

9. Malik H., Mishra S. Feature selection using RapidMiner and classification through probabilistic neural network for fault diagnostics of power transformer //2014 Annual IEEE India Conference (INDICON). – IEEE, 2014. – C. 1-6.
10. Ranjani J., Sheela A., Meena K. P. Combination of NumPy, SciPy and Matplotlib/PyLab-a good alternative methodology to MATLAB-A Comparative analysis //2019 1st international conference on innovations in information and communication technology (ICIICT). – IEEE, 2019. – C. 1-5.
11. Spendla L. et al. Concept of predictive maintenance of production systems in accordance with industry 4.0 //2017 IEEE 15Th International symposium on applied machine intelligence and informatics (SAMI). – IEEE, 2017. – C. 000405-000410.
12. Truong H. L. Integrated analytics for IIoT predictive maintenance using IoT big data cloud systems //2018 IEEE International Conference on Industrial Internet (ICII). – IEEE, 2018. – C. 109-118.
13. Abbasi T. et al. Development of predictive maintenance interface using multiple linear regression //2018 International Conference on Intelligent and Advanced System (ICIAS). – IEEE, 2018. – C. 1-5.
14. Ingole O. et al. Investigation of Different Regression Models For The Predictive Maintenance of Aircraft's Engine //2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT). – IEEE, 2022. – C. 1-6.
15. Vinh T. Q., Hoang P. M., Hoa T. Q. Machinery Monitoring and Maintenance System Using Decision Tree Model for Fault Prediction //2021 International Symposium on Electrical and Electronics Engineering (ISEE). – IEEE, 2021. – C. 36-41.
16. Wang H. et al. Research on Health Assessment and Fault Prediction Method of Aircraft Electromechanical System Based on Fault Tree //2023 International Conference on Mechatronics, IoT and Industrial Informatics (ICMIII). – IEEE, 2023. – C. 437-441.
17. Francis F., Mohan M. Arima model based real time trend analysis for predictive maintenance //2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). – IEEE, 2019. – C. 735-739.

18. Wang Y. et al. Software Usage Prediction Based on Hybrid LSTM-ARIMA Algorithm //2023 IEEE 5th International Conference on Power, Intelligent Computing and Systems (ICPICS). – IEEE, 2023. – C. 699-703.
19. Yang Y., Zheng H., Zhang R. Prediction and analysis of aircraft failure rate based on SARIMA model //2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA). – IEEE, 2017. – C. 567-571.
20. Shin J. H., Lee Y. S., Kim J. O. Fault prediction of wind turbine by using the SVM method //2014 International Conference on Information Science, Electronics and Electrical Engineering. – IEEE, 2014. – T. 3. – C. 1923-1926.
21. Yao N., Jia Y., Tao K. Rail weld defect prediction and related condition-based maintenance //Ieee Access. – 2020. – T. 8. – C. 103746-103758.
22. Zhang J. et al. A Predictive Maintenance Model of Equipment System Based on GRU Model //2022 International Conference on Virtual Reality, Human-Computer Interaction and Artificial Intelligence (VRHCIAI). – IEEE, 2022. – C. 100-105.
23. Yu Y. et al. Research of Fault Feature Extraction and Analysis Method Based on Aeroengine Fault Data //2020 Chinese Automation Congress (CAC). – IEEE, 2020. – C. 2960-2965.
24. Liao W., Li D. An improved prediction model for equipment performance degradation based on Fuzzy-Markov Chain //2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). – IEEE, 2015. – C. 6-10.
25. Zhang H., Chen M., Zhou D. Remaining useful life prediction for nonlinear degrading systems with maintenance //2017 Prognostics and System Health Management Conference (PHM-Harbin). – IEEE, 2017. – C. 1-5.
26. Liu C. R. et al. Improve production efficiency and predict machine tool status using Markov chain and hidden markov model //2018 8th International Conference on Computer Science and Information Technology (CSIT). – IEEE, 2018. – C. 276-281.
27. Reddy S. S. et al. Maintenance Enhancement of Smart Manufacturing Units In Industry 4.0 //2023 International Conference on Inventive Computation Technologies (ICICT). – IEEE, 2023. – C. 1488-1492.

28. Zhang Y. et al. Fault prediction of railway turnout systems based on improved sparse auto encoder and gated recurrent unit network //IEEE Transactions on Intelligent Transportation Systems. – 2021. – T. 23. – №. 8. – C. 12711-12723.
29. Liu L. et al. Densely Connected Fully Convolutional Auto-Encoder Based Slewing Bearing Degradation Trend Prediction Method //2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing). – IEEE, 2021. – C. 1-7.
30. Lu B. L. et al. A deep adversarial learning prognostics model for remaining useful life prediction of rolling bearing //IEEE Transactions on Artificial Intelligence. – 2021. – T. 2. – №. 4. – C. 329-340.
31. Fanan M. et al. Anomaly Detection for Hydroelectric Power Plants: a Machine Learning-based Approach //2023 IEEE 21st International Conference on Industrial Informatics (INDIN). – IEEE, 2023. – C. 1-6.
32. Mahmoud M. A. et al. A development methodology framework of smart manufacturing systems (Industry 4.0) //International Journal on Advanced Science, Engineering and Information Technology. – 2020. – T. 10. – №. 5. – C. 1927-1932.
33. Chang H. F. et al. A study of application of assessment modes in development of smart automatic production systems in Industry 4.0 //2016 International Conference on Fuzzy Theory and Its Applications (iFuzzy). – IEEE, 2016. – C. 1-8.
34. Chen B. et al. Smart factory of industry 4.0: Key technologies, application case, and challenges //Ieee Access. – 2017. – T. 6. – C. 6505-6519.
35. Cheng J. C. P. et al. Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms //Automation in Construction. – 2020. – T. 112. – C. 103087.
36. Schätz B. Platforms for cyber-physical systems-fractal operating system and integrated development environment for the physical world //2016 3rd International Workshop on Emerging Ideas and Trends in Engineering of Cyber-Physical Systems (EITEC). – IEEE, 2016. – C. 1-4.
37. Ribeiro L., Björkman M. Transitioning from standard automation solutions to cyber-physical production systems: an assessment of critical conceptual and technical challenges //IEEE systems journal. – 2017. – T. 12. – №. 4. – C. 3816-3827.

38. Zeb S. et al. Industrial digital twins at the nexus of nextG wireless networks and computational intelligence: A survey //Journal of Network and Computer Applications. – 2022. – C. 103309.
39. Vermesan O. et al. The next generation internet of things–hyperconnectivity and embedded intelligence at the edge //Next Generation Internet of Things. Distributed Intelligence at the Edge and Human Machine-to-Machine Cooperation. – 2018.
40. Çınar Z. M. et al. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0 //Sustainability. – 2020. – T. 12. – №. 19. – C. 8211.
41. Das S. et al. Essential steps in prognostic health management //2011 IEEE Conference on Prognostics and Health Management. – IEEE, 2011. – C. 1-9.
42. Zhang Y., Wang S., Han X. Research on decision-making process of condition-based maintenance //2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE). – IEEE, 2013. – C. 1437-1440.
43. Rawat M., Lad B. K. Condition based optimal maintenance strategy for multi-component system //2013 IEEE International Conference on Industrial Engineering and Engineering Management. – IEEE, 2013. – C. 1437-1441.
44. Rahhal J. S., Abualnadi D. IOT based predictive maintenance using LSTM RNN estimator //2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE). – IEEE, 2020. – C. 1-5.
45. Kanawaday A., Sane A. Machine learning for predictive maintenance of industrial machines using IoT sensor data //2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS). – IEEE, 2017. – C. 87-90.
46. Ruschel E., Santos E. A. P., Loures E. F. R. Mining shop-floor data for preventive maintenance management: integrating probabilistic and predictive models //Procedia Manufacturing. – 2017. – T. 11. – C. 1127-1134.
47. Lindstedt P., Sudakowski T., Rokicki E. Machine Diagnosis Based on Amplitude-Phase Characteristics, Determined from the Experimental Amplitude Spectrum and the Calculated Phase Spectrum //2020 International Conference Mechatronic Systems and Materials (MSM). – IEEE, 2020. – C. 1-4.

48. Patil S. S., Gaikwad J. A. Vibration analysis of electrical rotating machines using FFT: A method of predictive maintenance //2013 fourth international conference on computing, communications and networking technologies (ICCCNT). – IEEE, 2013. – C. 1-6.
49. Novoa C. G., Berríos G. A. G., Söderberg R. A. Predictive maintenance for motors based on vibration analysis with compact rio //2017 IEEE Central America and Panama Student Conference (CONESCAPAN). – IEEE, 2017. – C. 1-6.
50. Khademi A., Raji F., Sadeghi M. IoT enabled vibration monitoring toward smart maintenance //2019 3rd International Conference on Internet of Things and Applications (IoT). – IEEE, 2019. – C. 1-6.
51. Cahyono B., Arifianto I. Thermal condition assessment of power transformer //2009 IEEE 9th International Conference on the Properties and Applications of Dielectric Materials. – IEEE, 2009. – C. 60-62.
52. Yong Y. C. Research on Maintenance Method of Typical Aircraft Electronic Equipment Based on Thermal Analysis Technology //2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID). – IEEE, 2021. – C. 454-458.
53. Anoop G. L. et al. Thermal analysis of squirrel cage Induction Motor //2020 International Conference on Power Electronics and Renewable Energy Applications (PEREA). – IEEE, 2020. – C. 1-6.
54. Tareen S., Herzau J., Tianshu W. Predictive maintenance oriented pattern recognition system based on ultrasound data analysis //2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI). – IEEE, 2019. – C. 1208-1214.
55. Ma C. et al. Ultrasonic Waveguides Detection-based approach to locate defect on workpiece //2010 Chinese Control and Decision Conference. – IEEE, 2010. – C. 757-761.
56. Benyahia A., Benammar A., Guessoum A. Delamination Defects Localization in Carbon Fiber Composite Laminates Using Ultrasonic Signal Processing

//2020 6th International Symposium on New and Renewable Energy (SIENR). – IEEE, 2021. – C. 1-5.

57. Tzoumpas A., Anastasopoulos G., Kormentzas G. The importance of mixture of Thermography, Ultrasound Diagnostics & Power Analysis for energy audit and sustainable maintenance in Agricultural Industry //2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC). – IEEE, 2018. – C. 1307-1312.

58. Yang H. et al. Leak detection of gas transport pipelines based on wigner distribution //2011 International Symposium on Advanced Control of Industrial Processes (ADCONIP). – IEEE, 2011. – C. 258-261.

59. Karkulali P. et al. Leak detection in gas distribution pipelines using acoustic impact monitoring //IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society. – IEEE, 2016. – C. 412-416.

60. Jingwei G. et al. An integrated fault diagnosis method of gearboxes using oil analysis and vibration analysis //2007 8th International Conference on Electronic Measurement and Instruments. – IEEE, 2007. – C. 3-371-3-374.

61. Hossain R. R. et al. Pattern analysis of oil diagnostic results with service age of power transformers //2017 IEEE Calcutta Conference (CALCON). – IEEE, 2017. – C. 298-303.

62. Koksal A., Ozdemir A. An Extended Parametric Analysis of Transformer Maintenance Plan Using Markov State Diagram //2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). – IEEE, 2018. – C. 1-6.

63. Zhang Y. et al. A new maintenance decision making model based on life cycle cost analysis for power transformers //2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE). – IEEE, 2016. – C. 1-4.

64. Mainkar M. S., Rathod R. K. Maintenance Management System Effective Tool to Progress and Reach Total Productive Maintenance & Six Sigma Business Development Strategies //2017 International Conference on Industrial Engineering, Management Science and Application (ICIMSA). – IEEE, 2017. – C. 1-4.

65. Hermann M., Pentek T., Otto B. Design principles for industrie 4.0 scenarios //2016 49th Hawaii international conference on system sciences (HICSS). – IEEE, 2016. – C. 3928-3937.
66. Tantawi K. H., Sokolov A., Tantawi O. Advances in industrial robotics: From industry 3.0 automation to industry 4.0 collaboration //2019 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON). – IEEE, 2019. – C. 1-4.
67. Zhang L., Zhang J. A data-driven maintenance framework under imperfect inspections for deteriorating systems using multitask learning-based status prognostics //IEEE Access. – 2020. – T. 9. – C. 3616-3629.
68. Yan J. et al. Big-data-driven based intelligent prognostics scheme in industry 4.0 environment //2017 Prognostics and System Health Management Conference (PHM-Harbin). – IEEE, 2017. – C. 1-5.
69. Ruschel E., Santos E. A. P., Loures E. F. R. Establishment of maintenance inspection intervals: an application of process mining techniques in manufacturing //Journal of Intelligent Manufacturing. – 2020. – T. 31. – №. 1. – C. 53-72.
70. Singh G. K. et al. Induction machine drive condition monitoring and diagnostic research—a survey //Electric Power Systems Research. – 2003. – T. 64. – №. 2. – C. 145-158.
71. Luo X. et al. On research progress and development trend for motion control problems of industrial robots //2018 8th International Conference on Manufacturing Science and Engineering (ICMSE 2018). – Atlantis Press, 2018. – C. 241-246.
72. ZHANG L. I. N., WANG C. L. Research on Fault Diagnosis Method of Industrial Robots Based on Case-Based Reasoning //DEStech Transactions on Engineering and Technology Research. – 2018.
73. Han M., Shen L. H. Case-based reasoning based on FCM and neural network //Control and Decision. – 2012. – T. 27. – №. 9. – C. 1421-1424.
74. Chunling L., Zhichun M. Interval optimization of the industrial process based-on the case-based reasoning and its application //Proceedings of the 31st Chinese Control Conference. – IEEE, 2012. – C. 2234-2237.

75. Borgi T. et al. Data analytics for predictive maintenance of industrial robots //2017 International Conference on Advanced Systems and Electric Technologies (IC_ASET). – IEEE, 2017. – C. 412-417.
76. Li Z., He Q. Prediction of railcar remaining useful life by multiple data source fusion //IEEE Transactions on Intelligent Transportation Systems. – 2015. – T. 16. – №. 4. – C. 2226-2235.
77. Okoh C. et al. Overview of remaining useful life prediction techniques in through-life engineering services //Procedia Cirp. – 2014. – T. 16. – C. 158-163.
78. Sayyad S. et al. Data-driven remaining useful life estimation for milling process: sensors, algorithms, datasets, and future directions //IEEE Access. – 2021. – T. 9. – C. 110255-110286.
79. Bodo R., Bertocco M., Bianchi A. Fault classification driven by maintenance management for smart maintenance applications //2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT. – IEEE, 2020. – C. 27-32.
80. Dsouza J., Velan S. Preventive Maintenance for Fault Detection in Transfer Nodes using Machine Learning //2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). – IEEE, 2019. – C. 401-404.
81. Saha D. K., Ahmed S., Shaurov M. S. Different Machine Maintenance Techniques of Rotary Machine and Their Future Scopes: A Review //2019 4th International Conference on Electrical Information and Communication Technology (EICT). – IEEE, 2019. – C. 1-6.
82. Бурнаев Е. В. Алгоритмические основы предиктивной аналитики в задачах индустриального проектирования //Информационные процессы. – 2019. – Т. 19. – №. 3. – С. 249-270.
83. Zaytsev A., Burnaev E. Minimax approach to variable fidelity data interpolation //Artificial Intelligence and Statistics. – PMLR, 2017. – C. 652-661.
84. Zaytsev A., Burnaev E. Large scale variable fidelity surrogate modeling //Annals of Mathematics and Artificial Intelligence. – 2017. – Т. 81. – С. 167-186.

85. Kang Z., Catal C., Tekinerdogan B. Remaining useful life (RUL) prediction of equipment in production lines using artificial neural networks //Sensors. – 2021. – T. 21. – №. 3. – C. 932.
86. Silvestrin L. P., Hoogendoorn M., Koole G. A Comparative Study of State-of-the-Art Machine Learning Algorithms for Predictive Maintenance //SSCI. – 2019. – C. 760-767.
87. Chen J., Chen D., Liu G. Using temporal convolution network for remaining useful lifetime prediction //Engineering Reports. – 2021. – T. 3. – №. 3. – C. e12305.
88. Francis F., Mohan M. Arima model based real time trend analysis for predictive maintenance //2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). – IEEE, 2019. – C. 735-739.
89. Wang Y. et al. Software Usage Prediction Based on Hybrid LSTM-ARIMA Algorithm //2023 IEEE 5th International Conference on Power, Intelligent Computing and Systems (ICPICS). – IEEE, 2023. – C. 699-703.
90. Chamsukhee V. Time Series Forecasting of E-Databases Subscription Mahidol University Library with Exponential Smoothing, LSTM, and ARIMA Models //2020-5th International Conference on Information Technology (InCIT). – IEEE, 2020. – C. 202-207.
91. Sharp M. Observations on developing reliability information utilization in a manufacturing environment with case study: robotic arm manipulators //The International Journal of Advanced Manufacturing Technology. – 2019. – T. 102. – №. 9. – C. 3243-3264.
92. Hornung R. et al. Model-free robot anomaly detection //2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. – IEEE, 2014. – C. 3676-3683.
93. Bishop C. M., Nasrabadi N. M. Pattern recognition and machine learning. – New York: springer, 2006. – T. 4. – №. 4. – C. 738.
94. Nguyen-Tuong D., Peters J. Model learning for robot control: a survey //Cognitive processing. – 2011. – T. 12. – №. 4. – C. 319-340.

95. Pappachan B. K., Tjahjowidodo T. Parameter Prediction Using Machine Learning in Robot-Assisted Finishing Process //International Journal of Mechanical Engineering and Robotics Research. – 2020. – Т. 9. – №. 3.
96. Gopaluni R. B. et al. Modern machine learning tools for monitoring and control of industrial processes: A survey //IFAC-PapersOnLine. – 2020. – Т. 53. – №. 2. – С. 218-229.
97. Aliev K., Antonelli D. Proposal of a Monitoring System for Collaborative Robots to Predict Outages and to Assess Reliability Factors Exploiting Machine Learning //Applied Sciences. – 2021. – Т. 11. – №. 4. – С. 1621.
98. Ferreira L. et al. A comparison of AutoML tools for machine learning, deep learning and XGBoost //2021 International Joint Conference on Neural Networks (IJCNN). – IEEE, 2021. – С. 1-8.
99. Kocbek S., Gabrys B. Automated machine learning techniques in prognostics of railway track defects //2019 International Conference on Data Mining Workshops (ICDMW). – IEEE, 2019. – С. 777-784.
100. Tornede T. et al. AutoML for predictive maintenance: One tool to RUL them all //IoT Streams for Data-Driven Predictive Maintenance and IoT, Edge, and Mobile for Embedded Machine Learning: Second International Workshop, IoT Streams 2020, and First International Workshop, ITEM 2020, Co-located with ECML/PKDD 2020, Ghent, Belgium, September 14-18, 2020, Revised Selected Papers 2. – Springer International Publishing, 2020. – С. 106-118.
101. Schelter S. et al. On challenges in machine learning model management. – 2018.
102. MLOps for service management. Электронный ресурс. URL: <https://www.thedataops.org/how-to-use-mlops-for-service-management/>. [Дата обращения: 01.01.2022].
103. Garg S. et al. On continuous integration/continuous delivery for automated deployment of machine learning models using mlops //2021 IEEE fourth international conference on artificial intelligence and knowledge engineering (AIKE). – IEEE, 2021. – С. 25-28.

104. Ashmore R., Calinescu R., Paterson C. Assuring the machine learning lifecycle: Desiderata, methods, and challenges //ACM Computing Surveys (CSUR). – 2021. – Т. 54. – №. 5. – С. 1-39.
105. Oquab M. et al. Learning and transferring mid-level image representations using convolutional neural networks //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2014. – С. 1717-1724.
106. Sukhija S., Krishnan N. C., Kumar D. Supervised heterogeneous transfer learning using random forests //Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. – 2018. – С. 157-166.
107. Arief-Ang I. B., Hamilton M., Salim F. D. A scalable room occupancy prediction with transferable time series decomposition of CO2 sensor data //ACM Transactions on Sensor Networks (TOSN). – 2018. – Т. 14. – №. 3-4. – С. 1-28.
108. Larocque-Villiers J., Dumond P., Knox D. Automating predictive maintenance using state-based transfer learning and ensemble methods //2021 IEEE International Symposium on Robotic and Sensors Environments (ROSE). – IEEE, 2021. – С. 1-7.
109. Kwon N., Comuzzi M. Genetic algorithms for AutoML in process predictive monitoring //International Conference on Process Mining. – Cham : Springer Nature Switzerland, 2022. – С. 242-254.
110. Genetic algorithms in AutoML. Электронный ресурс. URL: <https://www.activeloop.ai/resources/glossary/genetic-algorithms-in-auto-ml/>.
111. Портал Towardsdatascience. Data-Driven Predictive Maintenance In a Nutshell. [Электронный ресурс]. Режим доступа: <https://towardsdatascience.com/data-driven-predictive-maintenance-in-a-nutshell-ccc65a13b998>. [Дата обращения 01.01.2022].
112. Мурзагулов Д. А., Замятин А. В. Адаптивные алгоритмы машинного обучения в управлении технологическими процессами //Автоматизация. Современные технологии. – 2018. – Т. 72. – №. 8. – С. 354-361.

113. Rostami M., Berahmand K., Forouzandeh S. A novel community detection based genetic algorithm for feature selection //Journal of Big Data. – 2021. – Т. 8. – №. 1. – С. 1-27.
114. Devi D. R., Sasikala S. Online Feature Selection (OFS) with Accelerated Bat Algorithm (ABA) and Ensemble Incremental Deep Multiple Layer Perceptron (EIDMLP) for big data streams //Journal of Big Data. – 2019. – Т. 6. – №. 1. – С. 1-20.
115. Kalia K., Gupta N. Analysis of hadoop MapReduce scheduling in heterogeneous environment //Ain Shams Engineering Journal. – 2021. – Т. 12. – №. 1. – С. 1101-1110.
116. Moradi P., Rostami M. A graph theoretic approach for unsupervised feature selection //Engineering Applications of Artificial Intelligence. – 2015. – Т. 44. – С. 33-45.
117. Cadenas J. M., Garrido M. C., MartíNez R. Feature subset selection filter-wrapper based on low quality data //Expert systems with applications. – 2013. – Т. 40. – №. 16. – С. 6241-6252.
118. Bonissone P. et al. A fuzzy random forest //International Journal of Approximate Reasoning. – 2010. – Т. 51. – №. 7. – С. 729-747.
119. Sun L., Xu J., Tian Y. Feature selection using rough entropy-based uncertainty measures in incomplete decision systems //Knowledge-Based Systems. – 2012. – Т. 36. – С. 206-216.
120. Tsamardinos I. et al. A greedy feature selection algorithm for Big Data of high dimensionality //Machine learning. – 2019. – Т. 108. – №. 2. – С. 149-202.
121. Renuka Devi D., Sasikala S. Accelerated simulated annealing and mutation operator feature selection method for big data //Int J Recent Technol Eng. – 2019. – Т. 8. – С. 910-6.
122. Feurer M., Hutter F. Hyperparameter optimization //Automated machine learning. – Springer, Cham, 2019. – С. 3-33.
123. Основные метрики задач классификации в машинном обучении [Электронный ресурс] // Webiomed.ru. URL: <https://webiomed.ru/blog/osnovnyye-metriki-zadach-klassifikatsii-v-mashinnom-obuchenii/>. [Дата обращения 01.01.2021].

124. Liu G. J. et al. Metric Learning Based Similarity Measure For Attribute Description Identification Of Energy Data //2020 International Conference on Machine Learning and Cybernetics (ICMLC). – IEEE, 2020. – C. 219-223.
125. Fazlic L. B. et al. A Machine Learning Approach for the Classification of Disease Risks in Time Series //2020 9th Mediterranean Conference on Embedded Computing (MECO). – IEEE, 2020. – C. 1-5.
126. Li W. M. et al. Similarity search over data stream using LPC-DTW //2007 International Conference on Machine Learning and Cybernetics. – IEEE, 2007. – T. 3. – C. 1631-1634.
127. Jeong T. Time-series Data Classification and Analysis associated with Machine Learning Algorithms for Cognitive Perception and Phenomenon //IEEE Access. – 2020. – T. 8. – C. 222417-222428.
128. Que Z., Xu Z. A data-driven health prognostics approach for steam turbines based on XGBoost and DTW //IEEE Access. – 2019. – T. 7. – C. 93131-93138.
129. Xu P. et al. Dynamic-time-warping-based measurement data alignment model for condition-based railroad track maintenance //IEEE Transactions on Intelligent Transportation Systems. – 2014. – T. 16. – №. 2. – C. 799-812.
130. Vartak M. et al. ModelDB: a system for machine learning model management //Proceedings of the Workshop on Human-In-the-Loop Data Analytics. – 2016. – C. 1-3.
131. Schelter S. et al. Automatically tracking metadata and provenance of machine learning experiments //Machine Learning Systems Workshop at NIPS. – 2017. – C. 27-29.
132. Baylor D. et al. Tfx: A tensorflow-based production-scale machine learning platform //Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2017. – C. 1387-1395.
133. Hynes N., Sculley D., Terry M. The data linter: Lightweight, automated sanity checking for ml data sets //NIPS MLSys Workshop. – 2017. – T. 1.

134. Cachada A. et al. Maintenance 4.0: Intelligent and predictive maintenance system architecture //2018 IEEE 23rd international conference on emerging technologies and factory automation (ETFa). – IEEE, 2018. – Т. 1. – С. 139-146.

135. Karim R. et al. Maintenance analytics—the new know in maintenance //IFAC-PapersOnLine. – 2016. – Т. 49. – №. 28. – С. 214-219.

136. Torim A. et al. Pattern Based Software Architecture for Predictive Maintenance //Symposium of the Norwegian AI Society. – Cham : Springer International Publishing, 2022. – С. 26-38.

137. Time series analysis with StumPy. [Электронный ресурс]. URL: <https://nabeelvalley.co.za/docs/time-series-data-analysis/time-series-analysis-with-stumpy/> [Дата обращения: 10.10.2022].

138. Aivaliotis P., Georgoulas K., Chrysolouris G. The use of Digital Twin for predictive maintenance in manufacturing //International Journal of Computer Integrated Manufacturing. – 2019. – Т. 32. – №. 11. – С. 1067-1080.

139. Centomo S., Dall’Ora N., Fummi F. The design of a digital-twin for predictive maintenance //2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFa). – IEEE, 2020. – Т. 1. – С. 1781-1788.

140. Liu Z. et al. Digital twin for predictive maintenance //NDE 4.0, Predictive Maintenance, Communication, and Energy Systems: The Digital Transformation of NDE. – 2023. – Т. 12489. – С. 27-37.

141. Aivaliotis P. et al. Methodology for enabling dynamic digital twins and virtual model evolution in industrial robotics-a predictive maintenance application //International Journal of Computer Integrated Manufacturing. – 2023. – С. 1-19.

142. Garan M., Tidriri K., Kovalenko I. A data-centric machine learning methodology: application on predictive maintenance of wind turbines //Energies. – 2022. – Т. 15. – №. 3. – С. 826.

143. What if federated machine learning. [Электронный ресурс]. URL: <https://bigdataschool.ru/blog/what-is-federated-machine-learning.html> [Дата обращения: 11.10.2022].

144. Ансамблиевые методы алгоритмов машинного обучения / Хабр. [Электронный ресурс]. URL: <https://habr.com/ru/articles/561732/> [Дата обращения: 12.10.2022].
145. Системная последовательность принятия решений. [Электронный ресурс]. URL: <https://helpiks.org/9-7888.html> [Дата обращения: 13.10.2022].
146. Data analysis and decision making. [Электронный ресурс]. URL: https://www.wu.ac.at/fileadmin/wu/d/i/ifr/Data_Analysis_and_Decision_Making.pdf [Дата обращения: 14.10.2022].
147. Tidriri K., Braydi A., Kazmi H. Data-driven Decision-Making Methodology for Prognostic and Health Management of Wind Turbines //2021 Australian & New Zealand Control Conference (ANZCC). – IEEE, 2021. – С. 104-109.
148. Carvalho T. P. et al. A systematic literature review of machine learning methods applied to predictive maintenance //Computers & Industrial Engineering. – 2019. – Т. 137. – С. 106024.
149. Culos A. et al. Integration of mechanistic immunological knowledge into a machine learning pipeline improves predictions //Nature machine intelligence. – 2020. – Т. 2. – №. 10. – С. 619-628.
150. D'Amour A. et al. Underspecification presents challenges for credibility in modern machine learning //The Journal of Machine Learning Research. – 2022. – Т. 23. – №. 1. – С. 10237-10297.
151. Kreuzberger D., Kühl N., Hirschl S. Machine learning operations (mlops): Overview, definition, and architecture //IEEE Access. – 2023.
152. Xin D. et al. Production machine learning pipelines: Empirical analysis and optimization opportunities //Proceedings of the 2021 International Conference on Management of Data. – 2021. – С. 2639-2652.
153. Нарке Н., Nelson С. Building machine learning pipelines. – O'Reilly Media, 2020.
154. Nikitin N. O. et al. Automated evolutionary approach for the design of composite machine learning pipelines //Future Generation Computer Systems. – 2022. – Т. 127. – С. 109-125.

155. Calabrese M. et al. SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0 //Information. – 2020. – T. 11. – №. 4. – C. 202.
156. Chen Y. W., Song Q., Hu X. Techniques for automated machine learning //ACM SIGKDD Explorations Newsletter. – 2021. – T. 22. – №. 2. – C. 35-50.
157. Stojanovic N., Jovic M. Continuous real-time anomaly detection in the flexible production: D2Lab-based use case //2018 IEEE International Conference on Big Data (Big Data). – IEEE, 2018. – C. 4213-4222.
158. Narayanan V., Bobba R. B. Learning based anomaly detection for industrial arm applications //Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy. – 2018. – C. 13-23.
159. Kayan H. et al. CASPER: Context-aware anomaly detection system for industrial robotic arms //2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops). – IEEE, 2023. – C. 282-284.
160. Yang K., Kpotufe S., Feamster N. An efficient one-class SVM for anomaly detection in the internet of things //arXiv preprint arXiv:2104.11146. – 2021.
161. Xu H. et al. Deep isolation forest for anomaly detection //IEEE Transactions on Knowledge and Data Engineering. – 2023.
162. Abdelrahman O., Keikhosrokiani P. Assembly line anomaly detection and root cause analysis using machine learning //IEEE Access. – 2020. – T. 8. – C. 189661-189672.
163. Fan L. et al. Comparative Study of Isolation Forest and LOF algorithm in anomaly detection of data mining //2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR). – IEEE, 2021. – C. 1-5.
164. Yin C. et al. Anomaly detection based on convolutional recurrent autoencoder for IoT time series //IEEE Transactions on Systems, Man, and Cybernetics: Systems. – 2020. – T. 52. – №. 1. – C. 112-122.

165. Tang T. W. et al. Anomaly detection neural network with dual auto-encoders GAN and its industrial inspection applications //Sensors. – 2020. – Т. 20. – №. 12. – С. 3336.

166. Goncharov A. et al. Automated anomalies detection in the work of industrial robots //IOP Conference Series: Materials Science and Engineering. – IOP Publishing, 2021. – Т. 1019. – №. 1. – С. 012095.

167. Чепкасов А. Ю., Писанкин А. С. Исследование зависимостей физических характеристик при работе промышленного манипулятора с использованием методов машинного обучения //Молодежь и современные информационные технологии: сборник трудов XVII Международной научно-практической конференции студентов, аспирантов и молодых учёных, 17-20 февраля 2020 г., г. Томск. – Томский политехнический университет, 2020. – С. 68-69.

Приложение 1 – Публикации по теме исследования

ВАК:

1) Гончаров А. С. Подход к формированию информационного обеспечения для процесса принятия решений по предиктивному техническому обслуживанию // Современные наукоемкие технологии. – 2024. – № 11. – С. 17-25.

2) Гончаров А. С., Савельев А. О., Писанкин А. С., Чепкасов А. Ю. Разработка алгоритма подбора модели анализа данных для прогнозирования работоспособности промышленных роботов // Моделирование, оптимизация и информационные технологии. – 2023. – Т. 11, – №4. – С. 1-12.

3) Гончаров А. С., Савельев А. О., Писанкин А. С., Чепкасов А. Ю., Джаякоди Д. Н. К. Подход к автоматическому прогнозированию состояния промышленных манипуляторов с применением методов машинного обучения // Доклады ТУСУР. – 2021. – Т. 24, № 1. – С. 48-54.

SCOPUS:

4) Goncharov, A. Automated anomalies detection in the work of industrial robots / A. Goncharov, A. Savelev, N. Krinitsyn, S. Mikhalevich // IOP Conference Series: Materials Science and Engineering. – 2019. – № 1019. – С. 012095.

РИНЦ:

5) Мангутова, Е. А. Обзор современных алгоритмов кластеризации данных / Е. А. Мангутова, А. С. Гончаров // Молодежь и современные информационные технологии: Сборник трудов XX Международной научно-практической конференции студентов, аспирантов и молодых ученых. – 2023. – № 1-1. – С. 242-243.

6) А.С. Гончаров. Сравнительный анализ методов машинного обучения для анализа данных об отказах оборудования / А.С. Гончаров // XVIII Международная научно-практическая конференция старшеклассников, студентов и аспирантов «Молодёжь и наука». – 2022. № 1-2. – С. 74-75.

7) А.С. Гончаров. Архитектура программного обеспечения управления прогнозными моделями «умного» производства / А.С. Гончаров // XXVII

Международная научно-техническая конференция студентов, аспирантов и молодых учёных «Научная сессия ТУСУР – 2022». – 2022. № 1-3. – С. 21-23.

8) А.С. Писанкин, А.Ю. Чепкасов, А.С. Гончаров. Разработка алгоритма подбора обученных моделей методами оценки подобия данных / А.С. Писанкин, А.Ю. Чепкасов, А.С. Гончаров // XXVII Международная научно-техническая конференция студентов, аспирантов и молодых учёных «Научная сессия ТУСУР – 2022». – 2022. № 1-3. – С. 119-121.

9) А.С. Писанкин, А.Ю. Чепкасов, А.С. Гончаров. Прогнозирование перегрева привода промышленного манипулятора с использованием методов интеллектуального анализа данных / А.С. Писанкин, А.Ю. Чепкасов, А.С. Гончаров // Международная научно-практическая конференция «Электронные средства и системы управления» (МНПК ЭСиСУ-2020). – 2020. № 1-2. – С. 10-12.

10) А.С. Гончаров. Разработка алгоритма автоматизированного поиска аномальных значений параметров промышленных манипуляторов / А.С. Гончаров // XX Всероссийская конференция молодых учёных по математическому моделированию и информационным технологиям: тезисы докладов. – 2019. № 1-1. – С. 57.

Приложение 2 – Результаты интеллектуальной деятельности



Рисунок 54 – Свидетельство о регистрации программы для ЭВМ «Диагностика аномальных состояний приводов промышленных манипуляторов»



Рисунок 55 – Свидетельство о регистрации программы для ЭВМ «Модуль адаптивного выбора прогнозной модели работоспособности промышленного манипулятора»

Приложение 3 – Диаграммы процесса принятия решений по ТОиР на основе анализа данных (в нотации IDEF0)

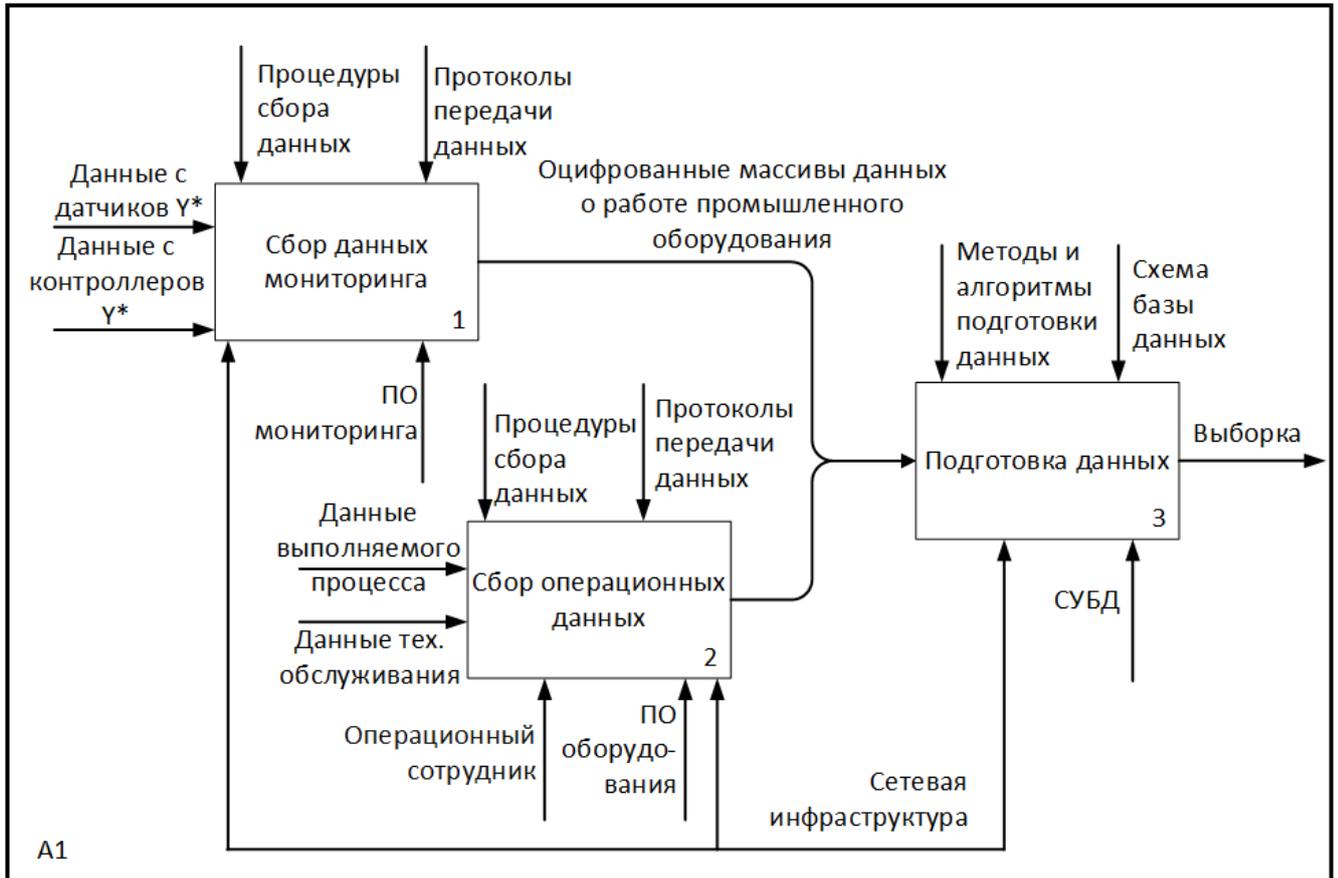


Рисунок 56 – Процесс сбора данных для формирования выборки (блок А1)

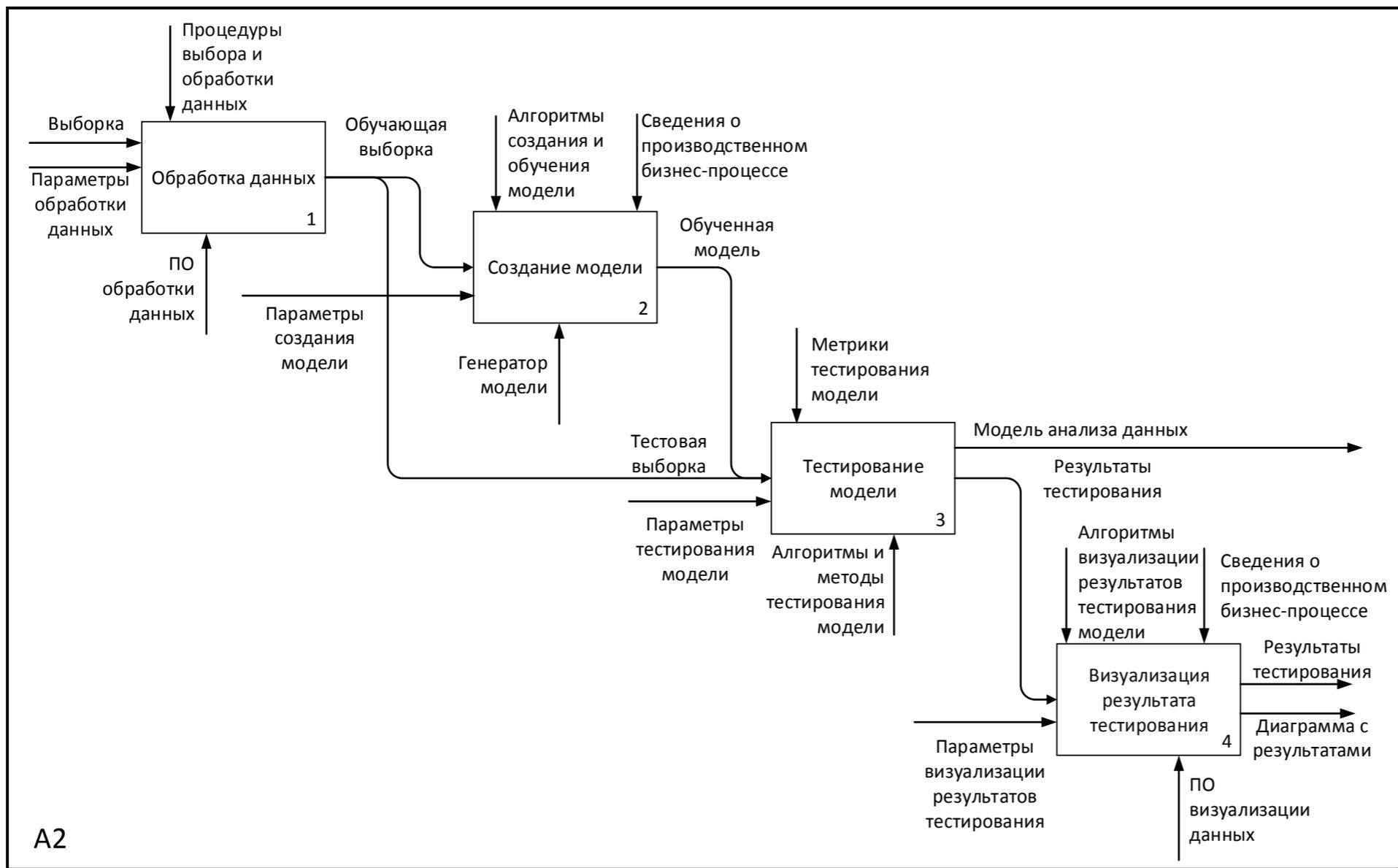


Рисунок 57 – Процесс создания и тестирования модели анализа данных (блок А2)

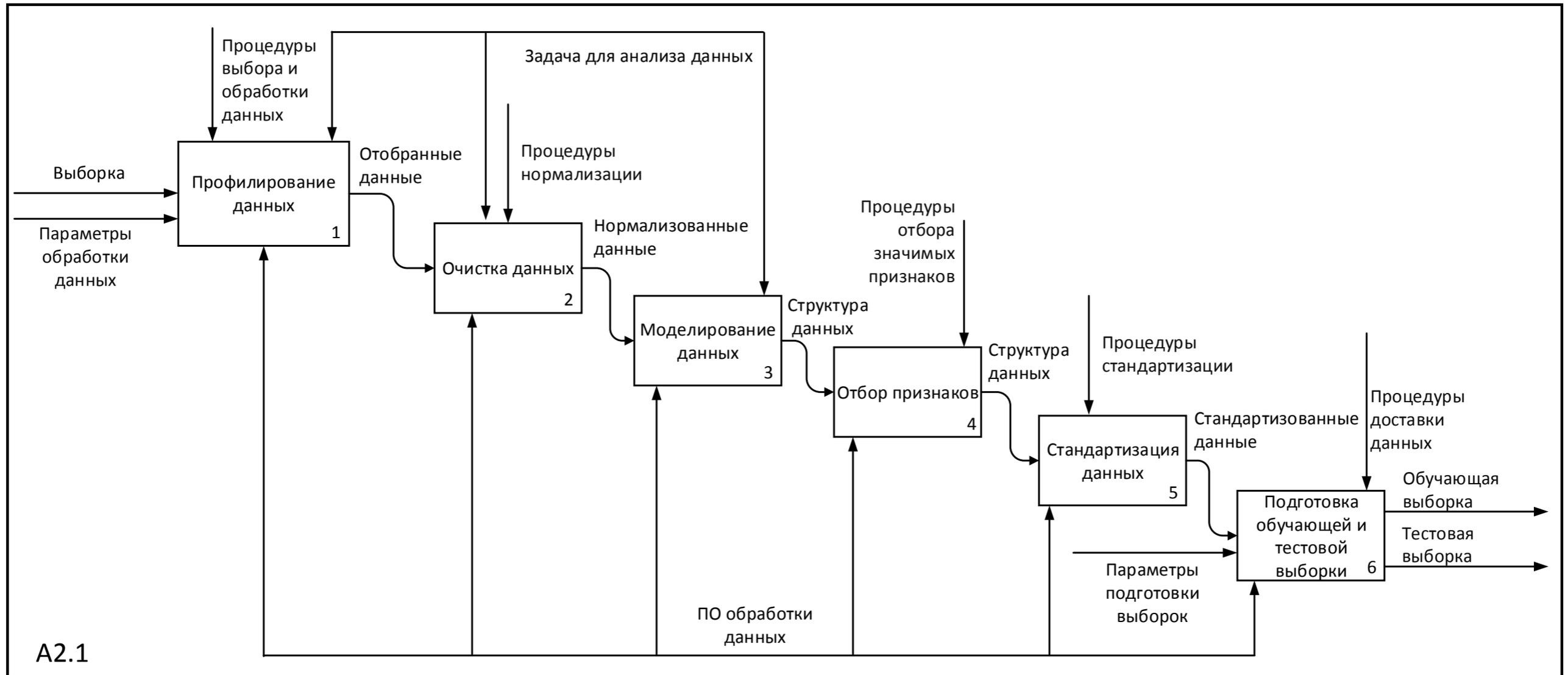


Рисунок 58 – Процесс обработки данных для анализа (блок A2.1)

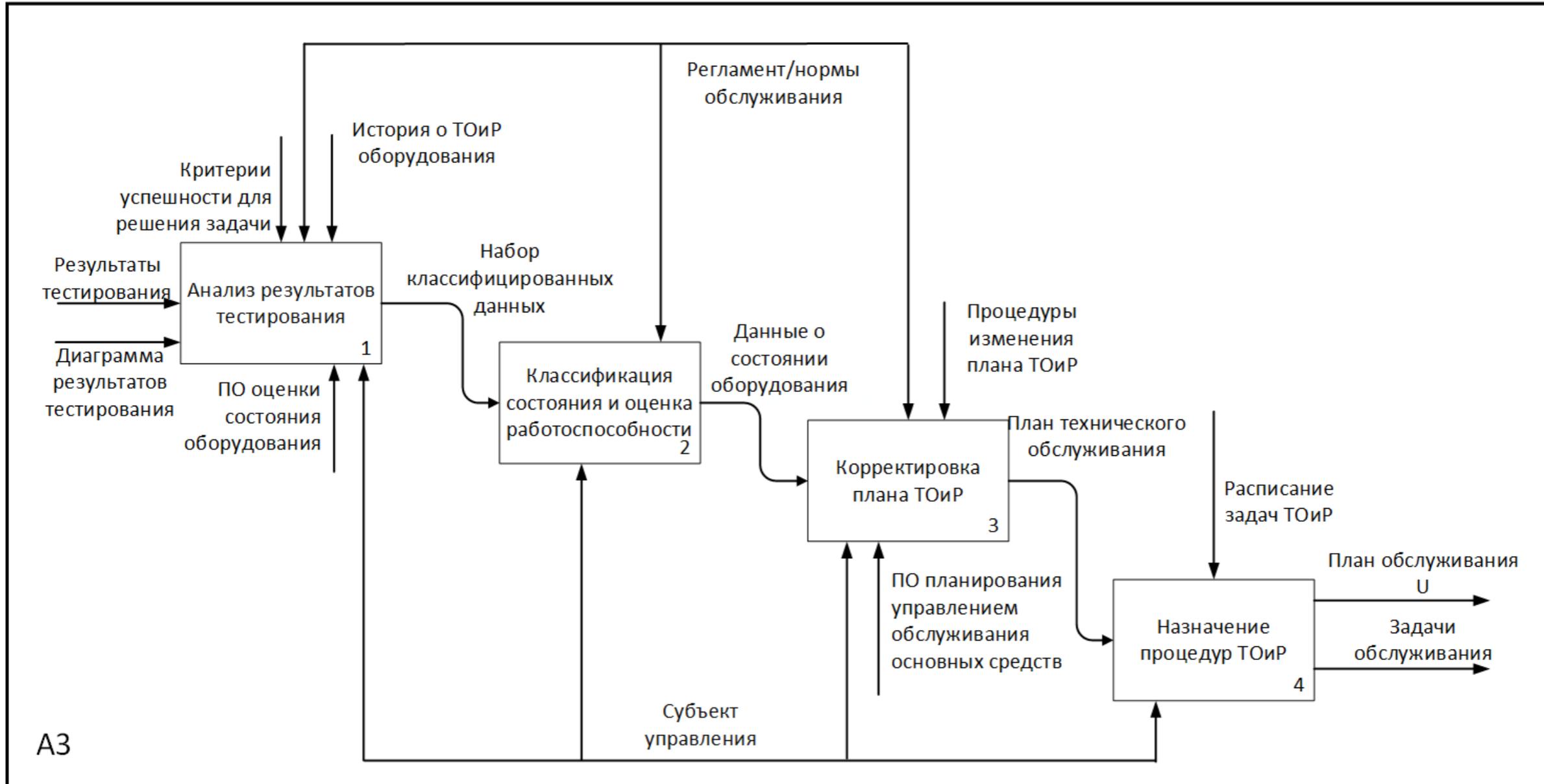


Рисунок 59 – Процесс корректировки плана ТОиР на основе анализа данных (блок А3)

Приложение 4 – Логическая схема базы данных для хранения исторических данных о техническом обслуживании промышленного оборудования

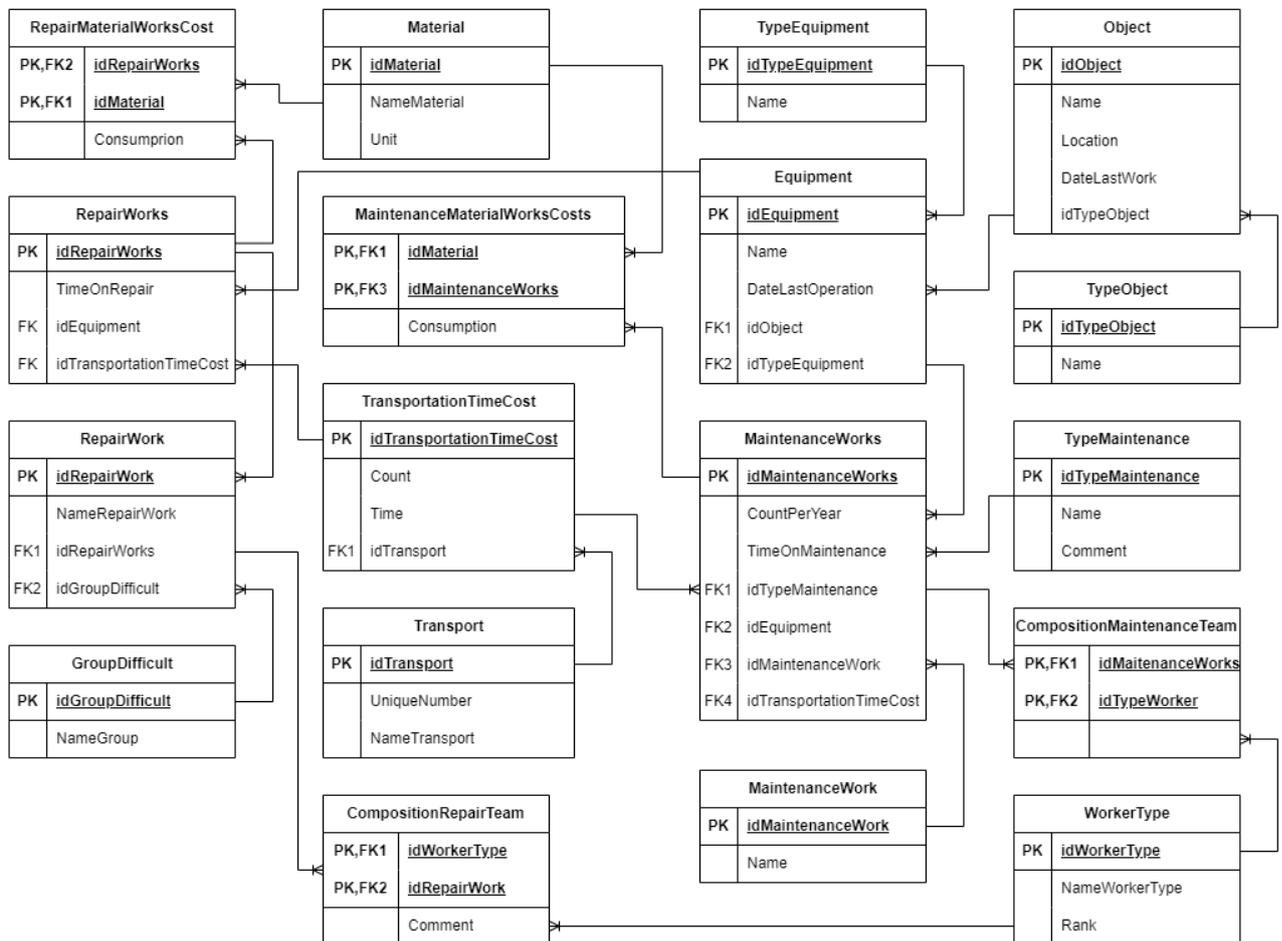


Рисунок 60 – Логическая схема базы данных для хранения информации о техническом обслуживании и ремонта

База данных для хранения информации о техническом обслуживании включает в себя следующие таблицы: Ремонтные работы, Стоимость ремонтных работ, Тип и категория ремонтных работ, оборотные средства, Транспортировка и её стоимость, Состав ремонтной бригады, Оснастка, Тип и стоимость технического обслуживания.

Приложение 5 – Алгоритм обучения и тестирования модели на наборах данных разнородного оборудования

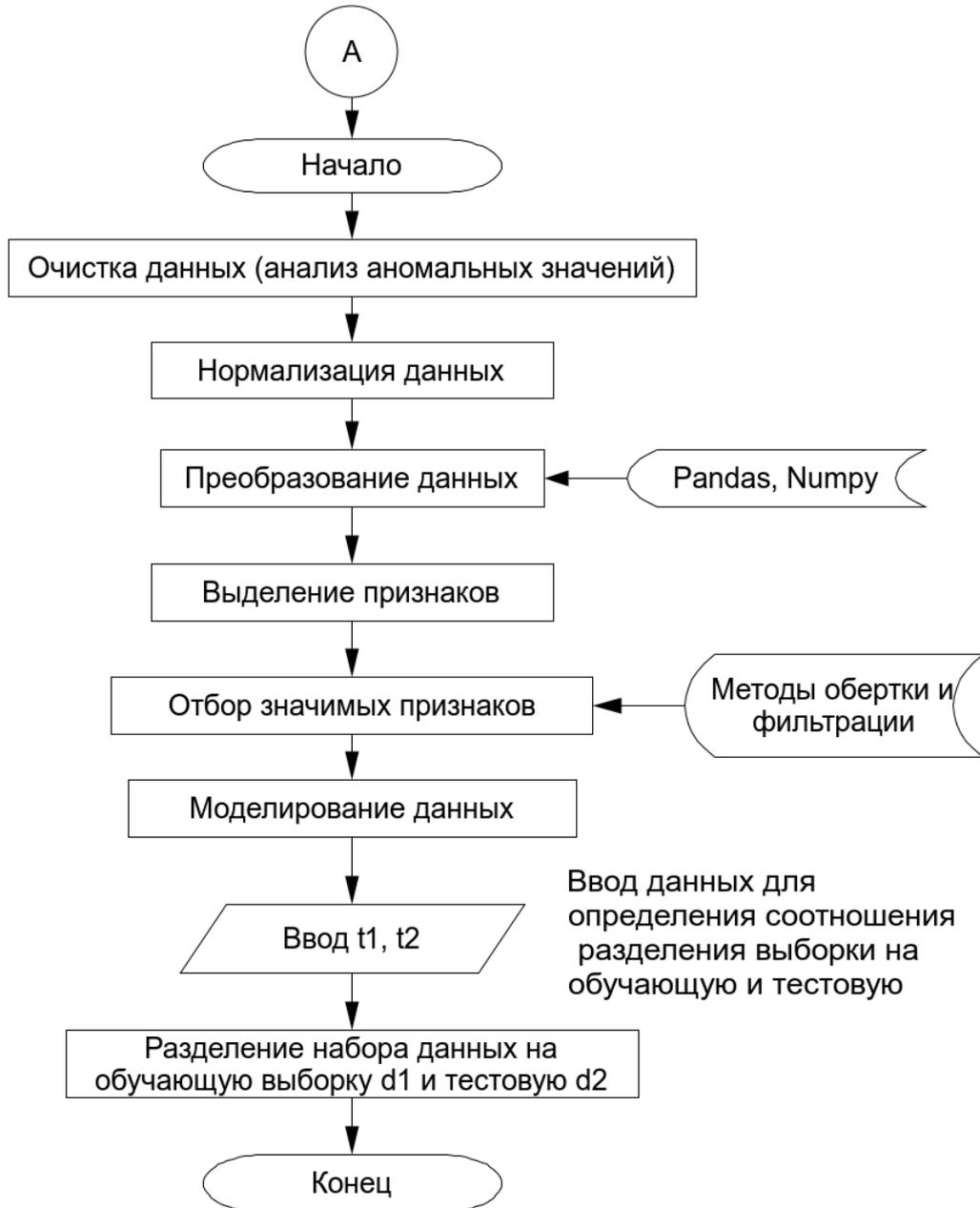


Рисунок 61 – Функция предобработки данных

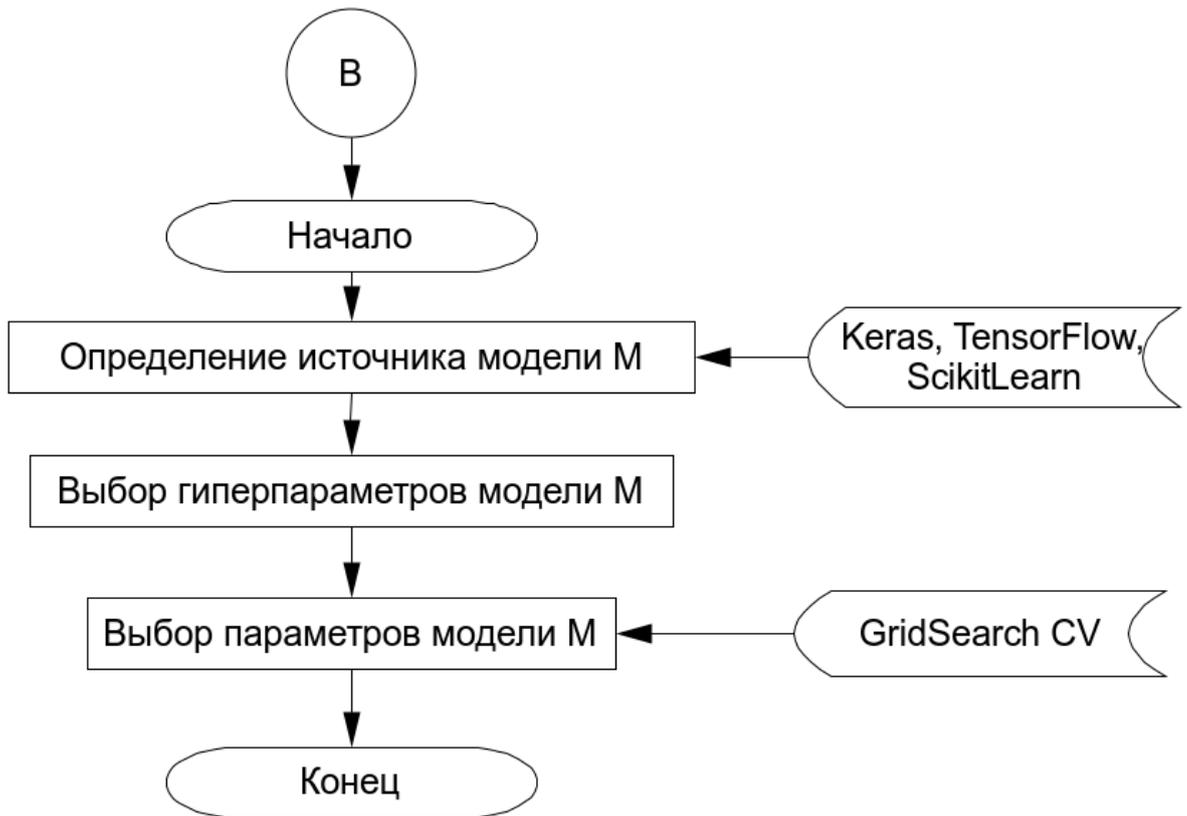


Рисунок 62 – Конструирования модели

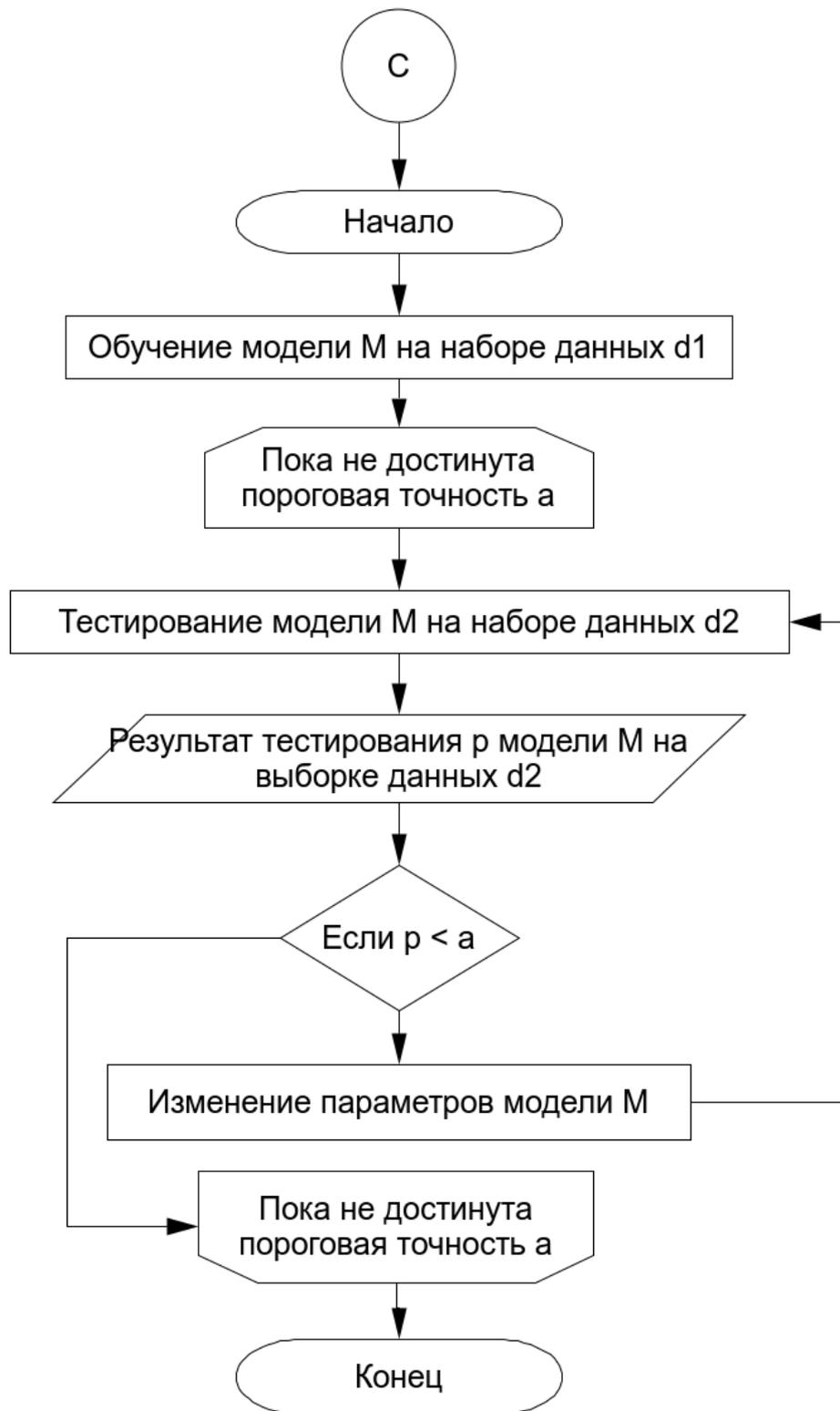


Рисунок 63 – Алгоритм обучения модели



Рисунок 64 – Тестирование модели

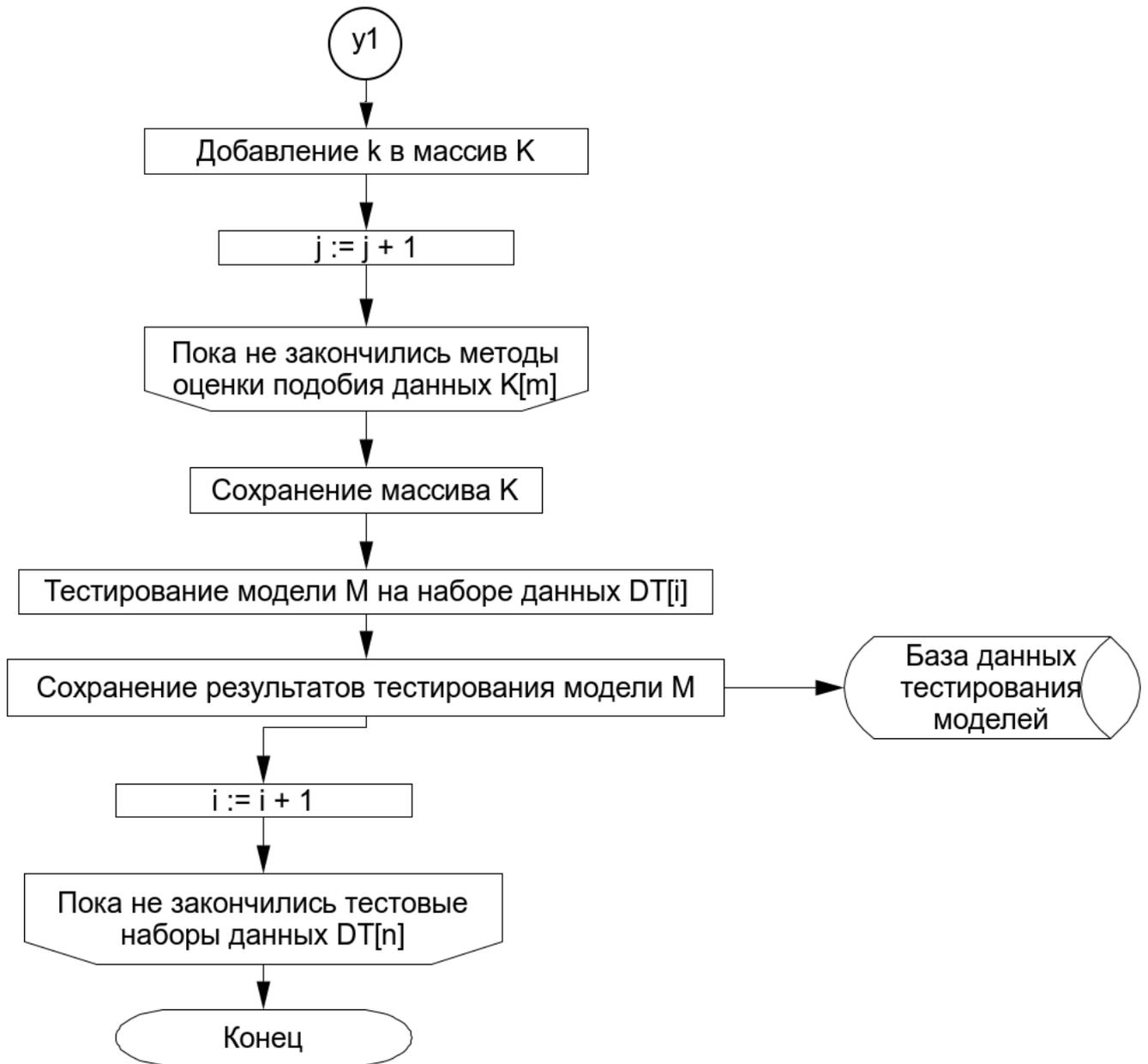


Рисунок 65 – Тестирование модели

Приложение 6 – Алгоритмы обучения и тестирования модели анализа данных

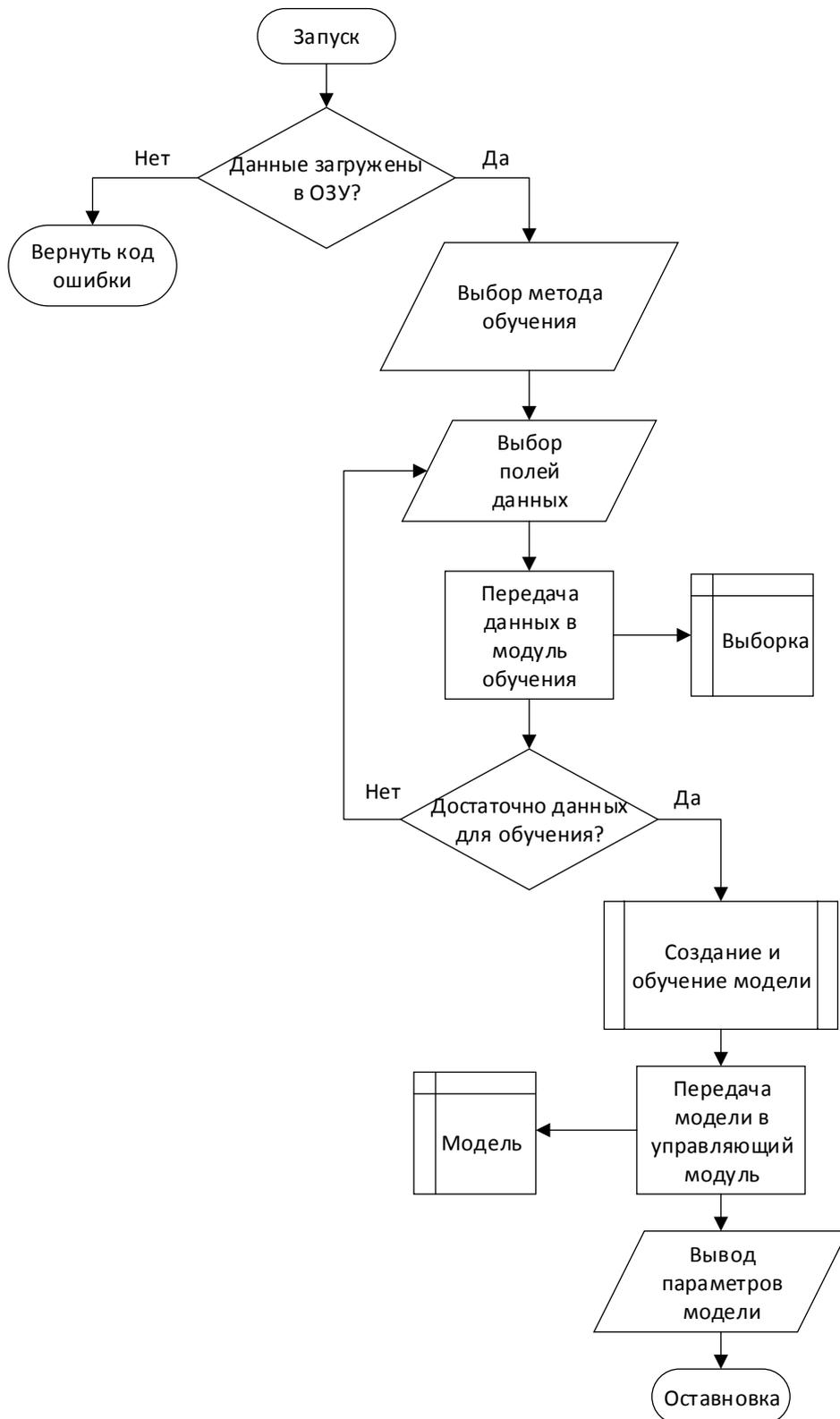


Рисунок 6б – Алгоритм обучения модели на данных из оперативной памяти

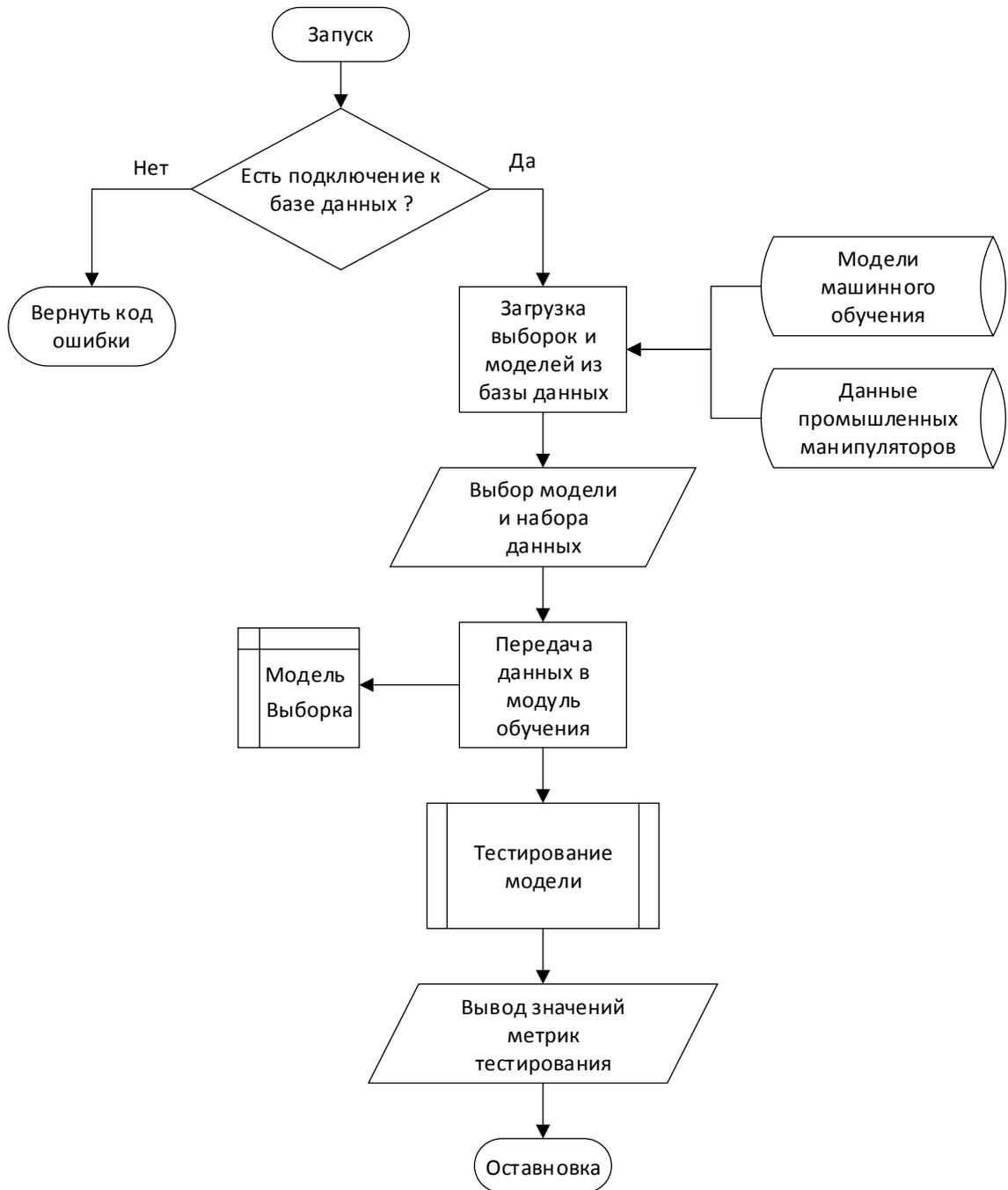


Рисунок 67 – Тестирование моделей на датасетах, загружаемых из базы данных

Приложение 7 – Структура файла для апробации алгоритмической модели

```

_DATE_,_NumMotor_,_CURR_ACT_,_VEL_AXIS_ACT_,_MOT_TEMP_,_TORQUE_AXIS_AC
T_,_VEL_AXIS_,
14:52:41,_1_,-1.25885,-1.00581,300,15.5715,70,
14:52:41,_2_,-3.07208,18.5585,304,30.5508,70,
14:52:41,_3_,0.349914,-8.60341,303,-1.70975,70,
14:52:41,_4_,-3.55746,-12.4741,303,-4.1036,70,
14:52:41,_5_,-8.00333,-4.37419,309,-9.03176,70,
14:52:41,_6_,9.6823,2.5266,309,6.89619,70,
14:52:42,_1_,4.62408,1.8692,300,-57.1795,70,
14:52:42,_2_,-1.98945,-0.358581,304,19.7877,70,
14:52:42,_3_,-4.87566,-2.40835,303,23.8115,70,
14:52:42,_4_,0.327512,-0.0305174,302,0.378005,70,
14:52:42,_5_,2.2198,0.106811,309,2.50679,70,
14:52:42,_6_,-8.91936,0.00254306,309,-6.35385,70,
14:52:43,_1_,1.49561,1.9544,300,-18.5,70,
14:52:43,_2_,-5.05323,-33.4689,304,50.2276,70,
14:52:43,_3_,9.4526,15.5996,303,-46.0979,70,
14:52:43,_4_,6.38711,22.2181,302,7.36835,70,
14:52:43,_5_,10.7677,7.60396,309,12.144,70,
14:52:43,_6_,-15.0932,-4.44794,309,-10.7332,70,
14:52:44,_1_,-4.44329,-34.5116,300,54.9454,70,
14:52:44,_2_,5.8228,-6.37309,304,-57.8625,70,
14:52:44,_3_,8.34886,11.0283,303,-40.7336,70,
14:52:44,_4_,3.74066,0.349681,303,4.31506,70,
14:52:44,_5_,-11.0626,-10.7295,309,-12.4757,70,
14:52:44,_6_,-6.8652,-0.84432,309,-4.89242,70,
14:52:45,_1_,5.15565,33.5375,301,-63.7277,70,
14:52:45,_2_,13.7441,6.24084,304,-135.941,70,
14:52:45,_3_,-1.6535,-10.8973,303,8.07887,70,
14:52:45,_4_,-2.99153,-0.450134,303,-3.45095,70,
14:52:45,_5_,11.5156,10.5362,309,12.985,70,
14:52:45,_6_,5.76848,0.750224,309,4.11152,70,
14:52:46,_1_,-1.70583,-1.7217,301,21.093,70,
14:52:46,_2_,15.7369,30.0344,304,-155.374,70,
14:52:46,_3_,-2.16751,-13.8524,303,10.5898,70,
14:52:46,_4_,-5.97427,-20.0132,302,-6.89248,70,
14:52:46,_5_,-10.3762,-6.82322,309,-11.7036,70,
14:52:46,_6_,13.4807,4.07918,309,9.5917,70,
14:52:47,_1_,-3.82548,-30.8584,301,47.2937,70,
14:52:47,_2_,0.840063,-3.69008,304,-8.35637,70,

```

Приложение 8 – Комбинации признаков для отбора значимых параметров

Таблица 12 – Наборы атрибутов для тестирования

№ опыта	Выбранные параметры
1	Сила тока
	Температура привода
	Момент силы
2	Сила тока
	Температура привода
	Измеренная скорость вращения
3	Сила тока
	Температура привода
	Момент силы
	Измеренная скорость вращения
4	Сила тока
	Температура привода
	Момент силы
	Измеренная скорость вращения
	Заданная скорость вращения
5	Температура привода
	Измеренная скорость вращения
6	Сила тока
	Температура привода

Таблица 13 – Результаты тестирования данных с разными наборами атрибутов

Модель	№ опыта	T, сек	Кол-во эпох	Точность, %
Линейная модель	1	14,3675	9	73,1
	2	25,4782	15	71,6437
	3	21,96	14	72,0875
	4	3,94	2	0
	5	22,1596	11	71,2125
	6	6,61	4	80,6562
Многослойный перцептрон	1	14	7	72,718
	2	9,55662	5	71,84375
	3	12,40	6	72,256
	4	4,678	2	0
	5	5,79	3	71,2624
	6	14,7603	8	81,15
Сверточная ИНС	1	13,69	7	69,6812
	2	29,81979	16	70,74375
	3	11,59	6	66,64
	4	4,48	2	0
	5	21,2335	12	68,7312
	6	26,95465	15	72,6062

Приложение 9 – Результаты нормализации датасетов

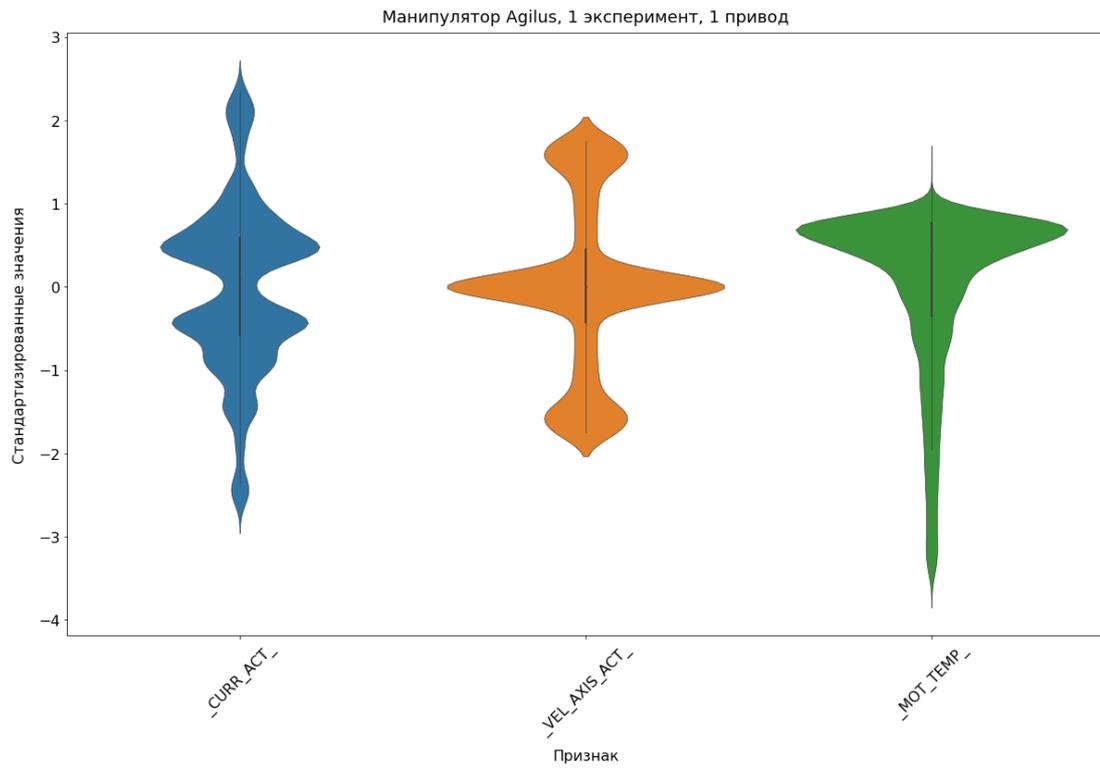


Рисунок 68 – Пример нормализации данных робота «agilus»

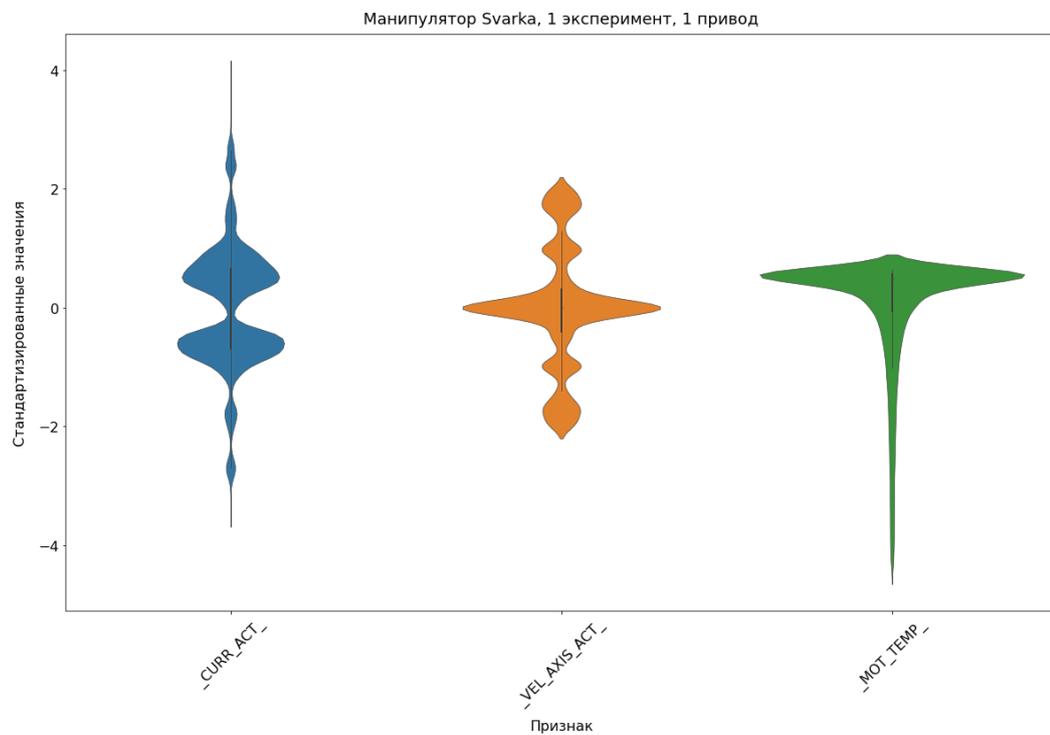


Рисунок 69 – Пример нормализации данных робота «svarka»

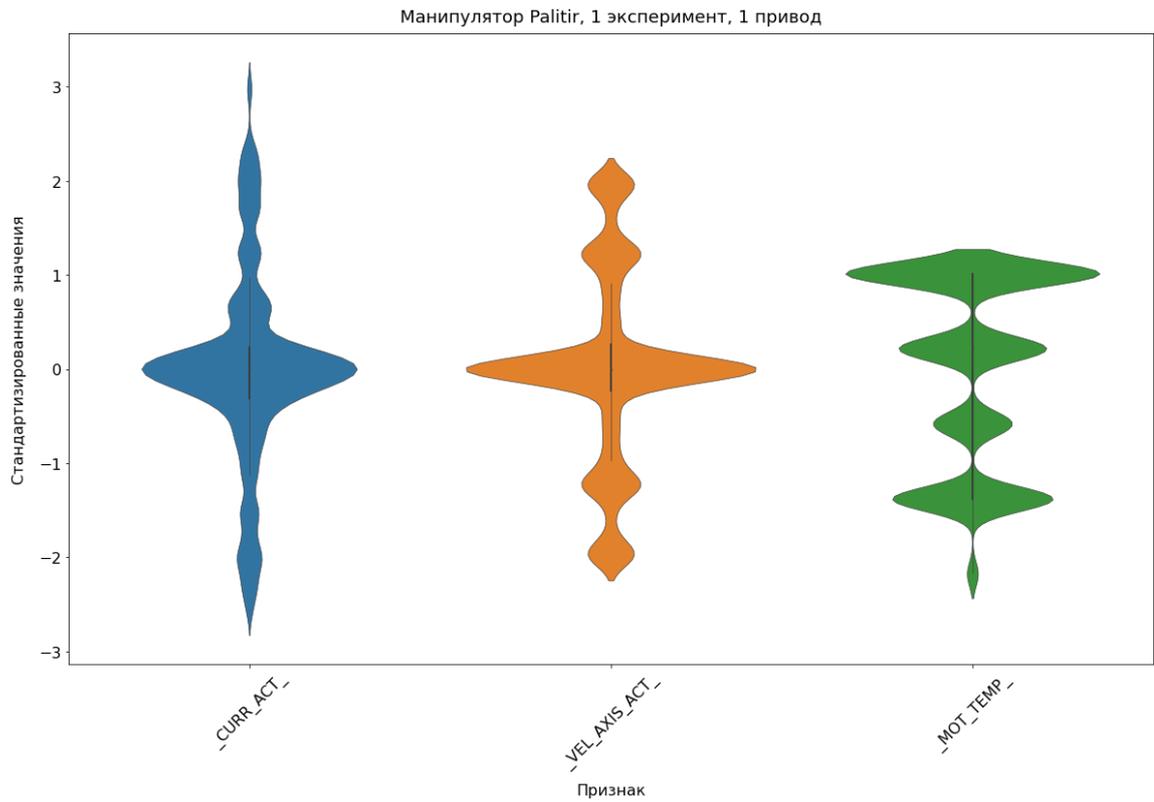


Рисунок 70 – Пример нормализации данных робота «palitir»

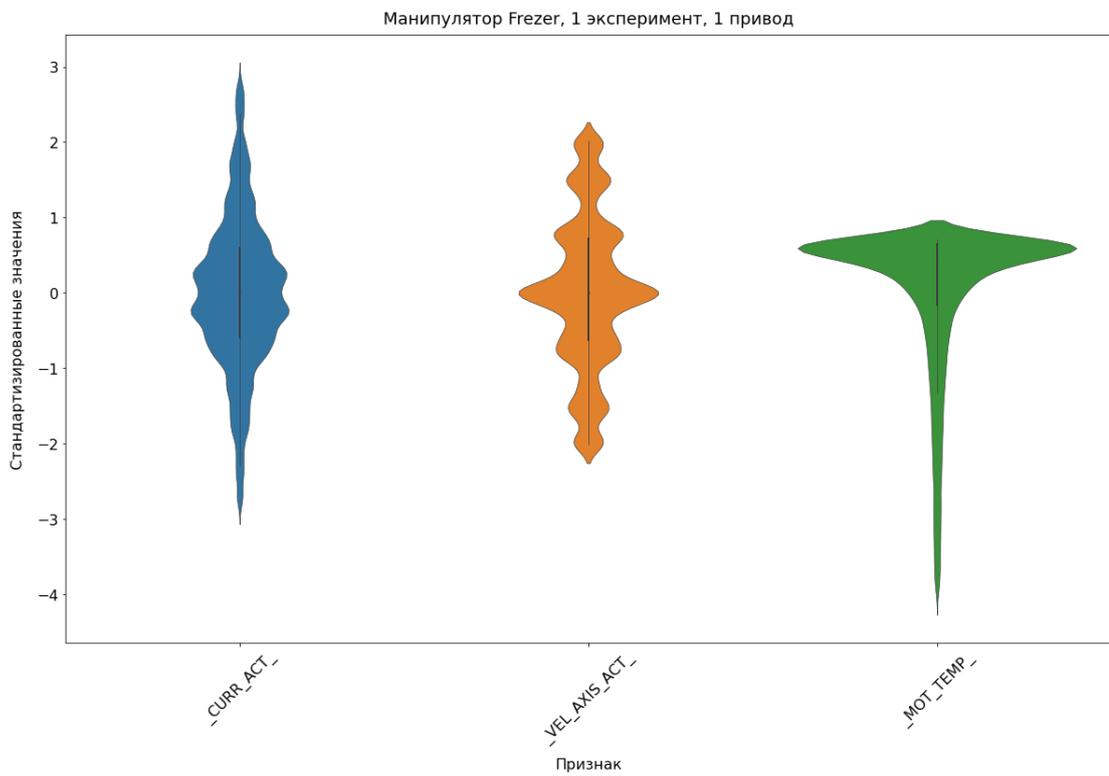


Рисунок 71 – Пример нормализации данных робота «frezer»

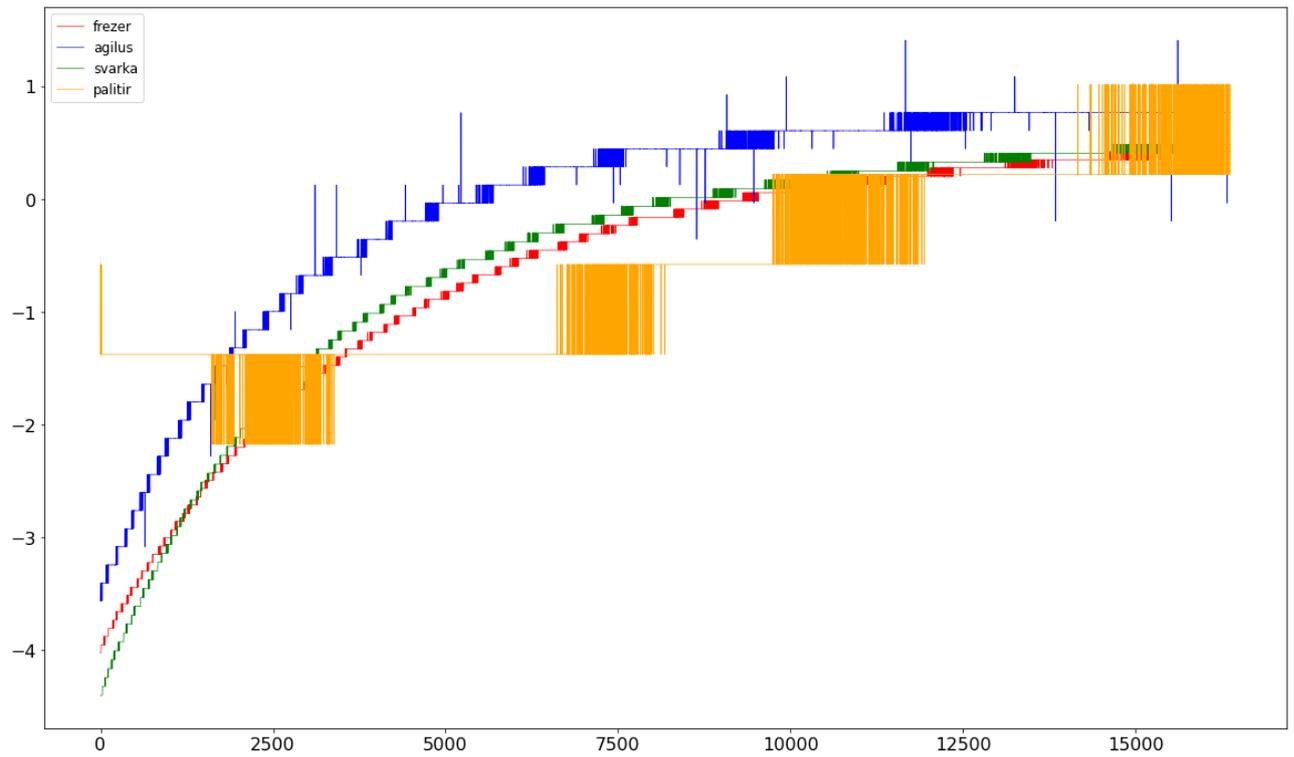


Рисунок 72 – Пример визуализации нормализованных данных 4-х промышленных роботов («agilus», «svarka», «frezer», «palitir»)

Приложение 10 – Графики зависимости точности моделей машинного обучения от сходства датасетов на примере задачи прогнозирования температуры привода

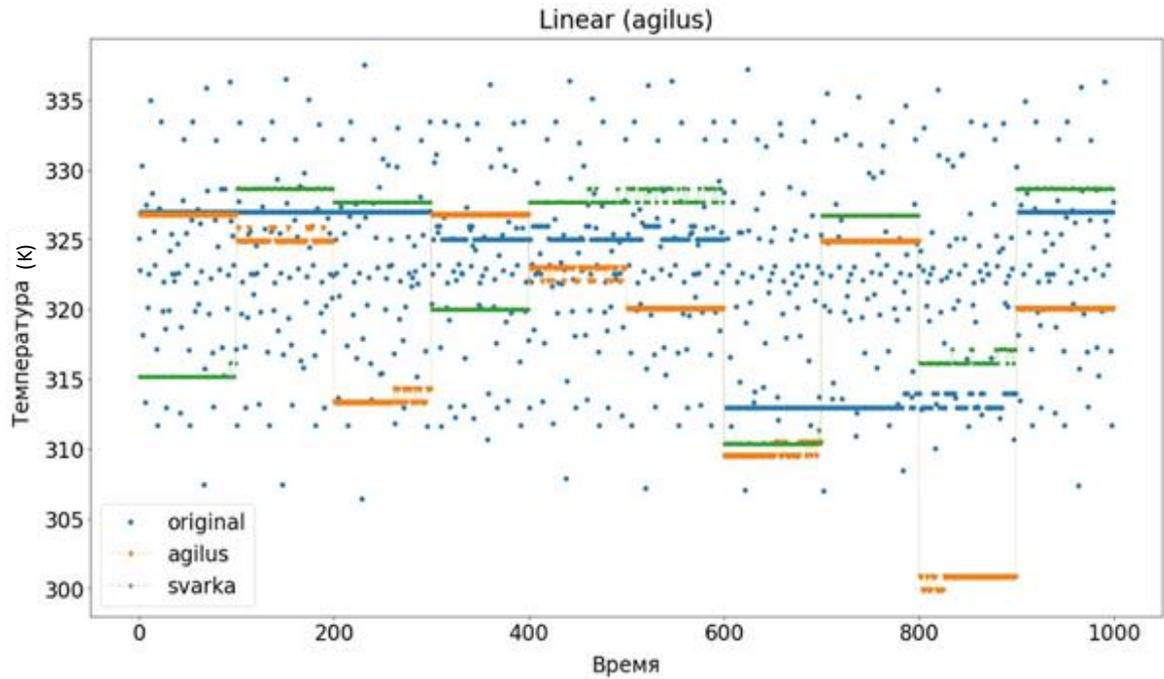


Рисунок 73 – Прогнозирование температуры привода линейной моделью, обученной на работе «agilus» и протестированной на работе «svarka»

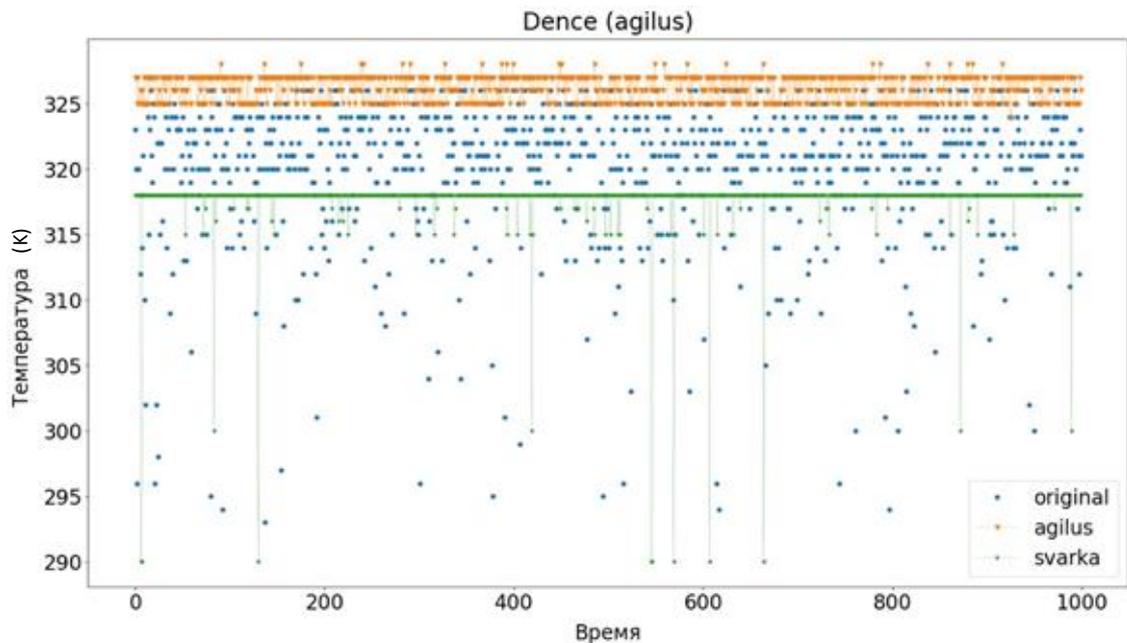


Рисунок 74 – Прогнозирование температуры привода моделью многослойный перцептрон, обученной на работе «agilus» и протестированной на работе «svarka»

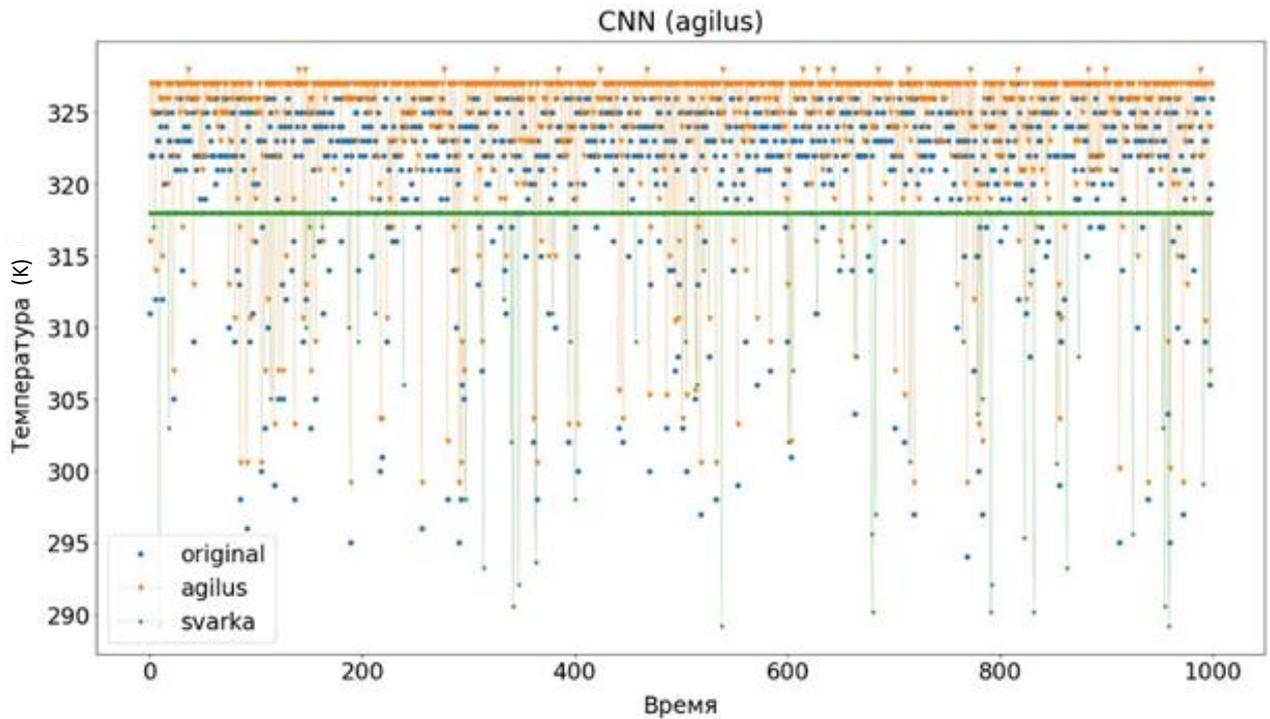


Рисунок 75 – Прогнозирование температуры привода моделью сверточной нейронной сети, обученной на работе «agilus» и протестированной на работе «svarka»

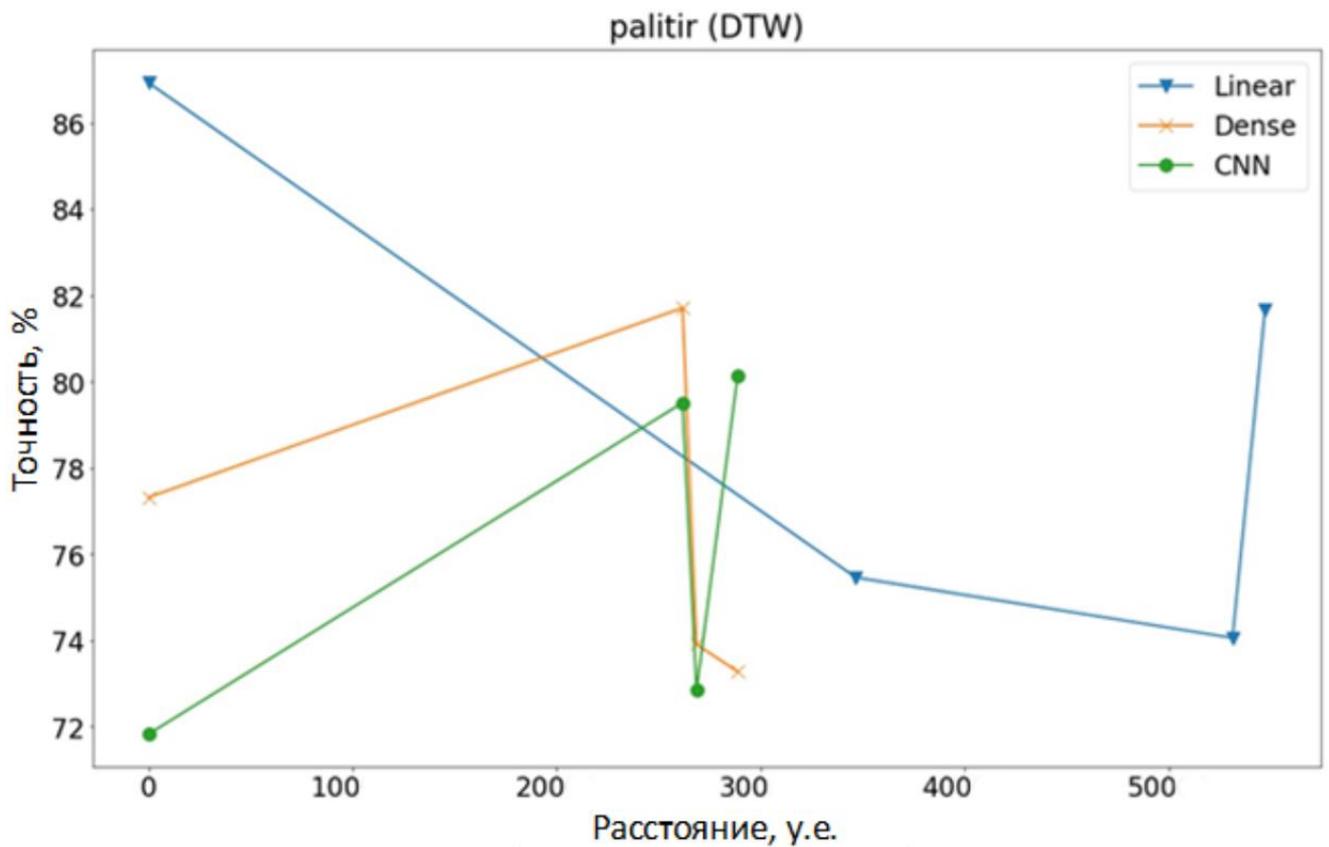


Рисунок 76 – Зависимость точности тестирования моделей от сходства датасетов (данные для обучения модели взяты в работа «palitir»)

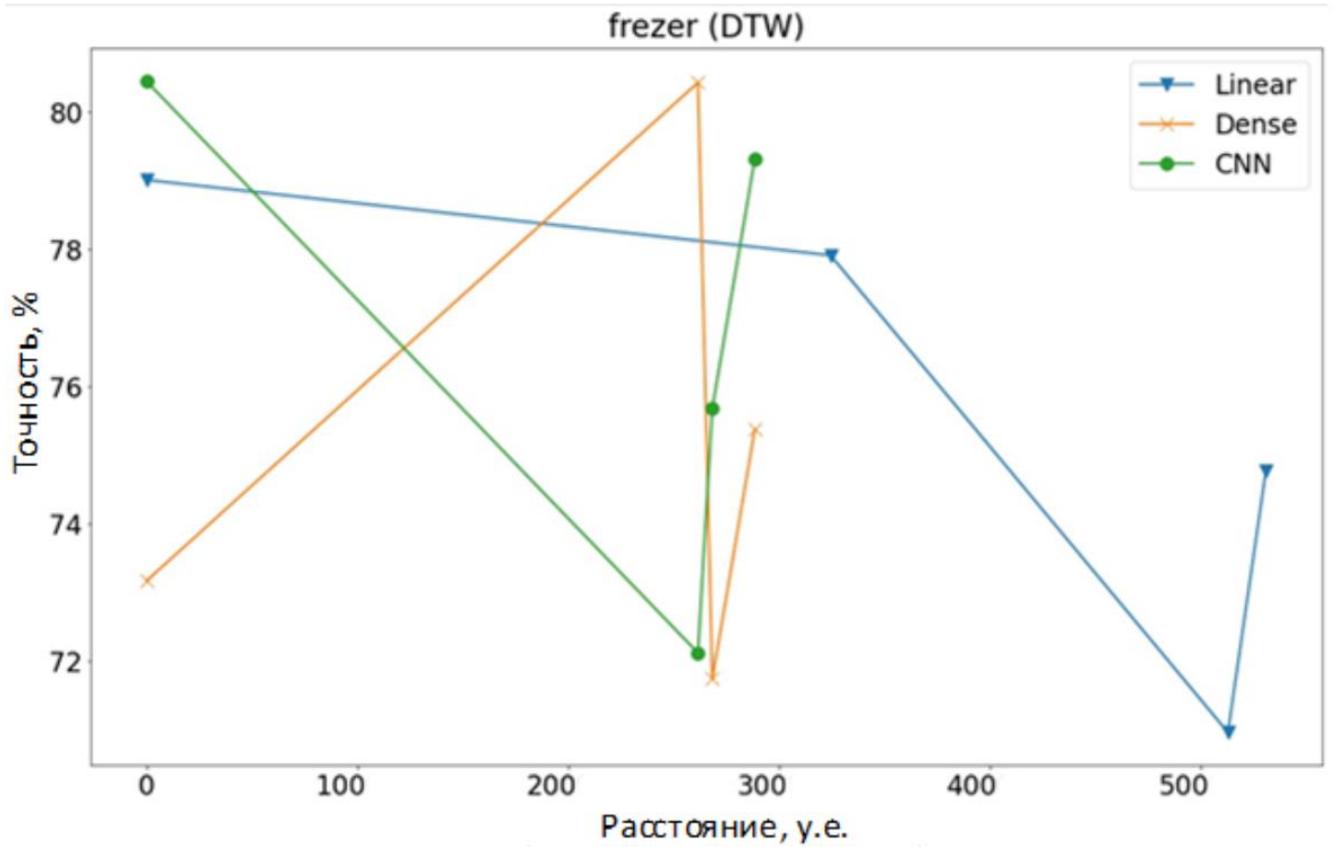


Рисунок 77 – Зависимость точности тестирования моделей от сходства датасетов (данные для обучения модели взяты в робота «frezer»)

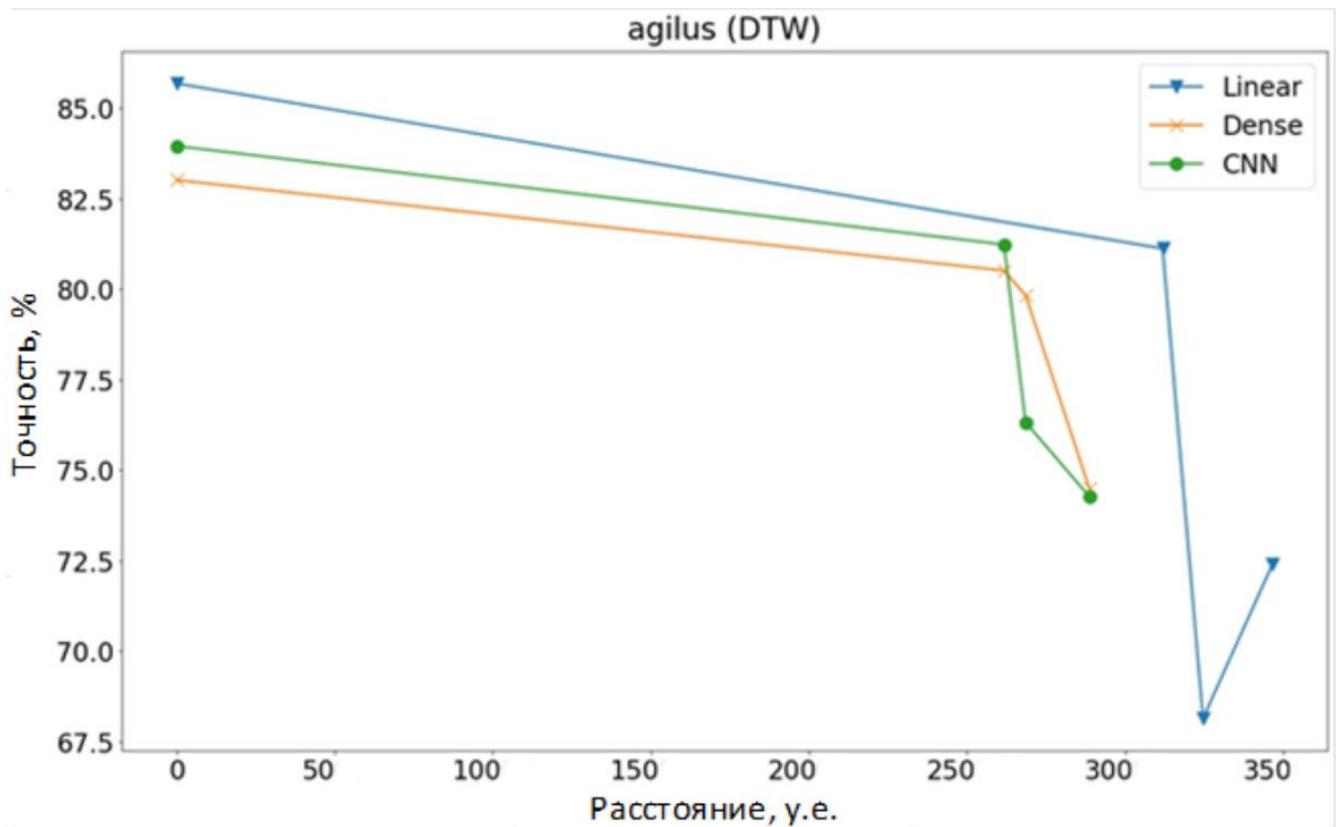


Рисунок 78 – Зависимость точности тестирования моделей от сходства датасетов (данные для обучения модели взяты в робота «agilus»)

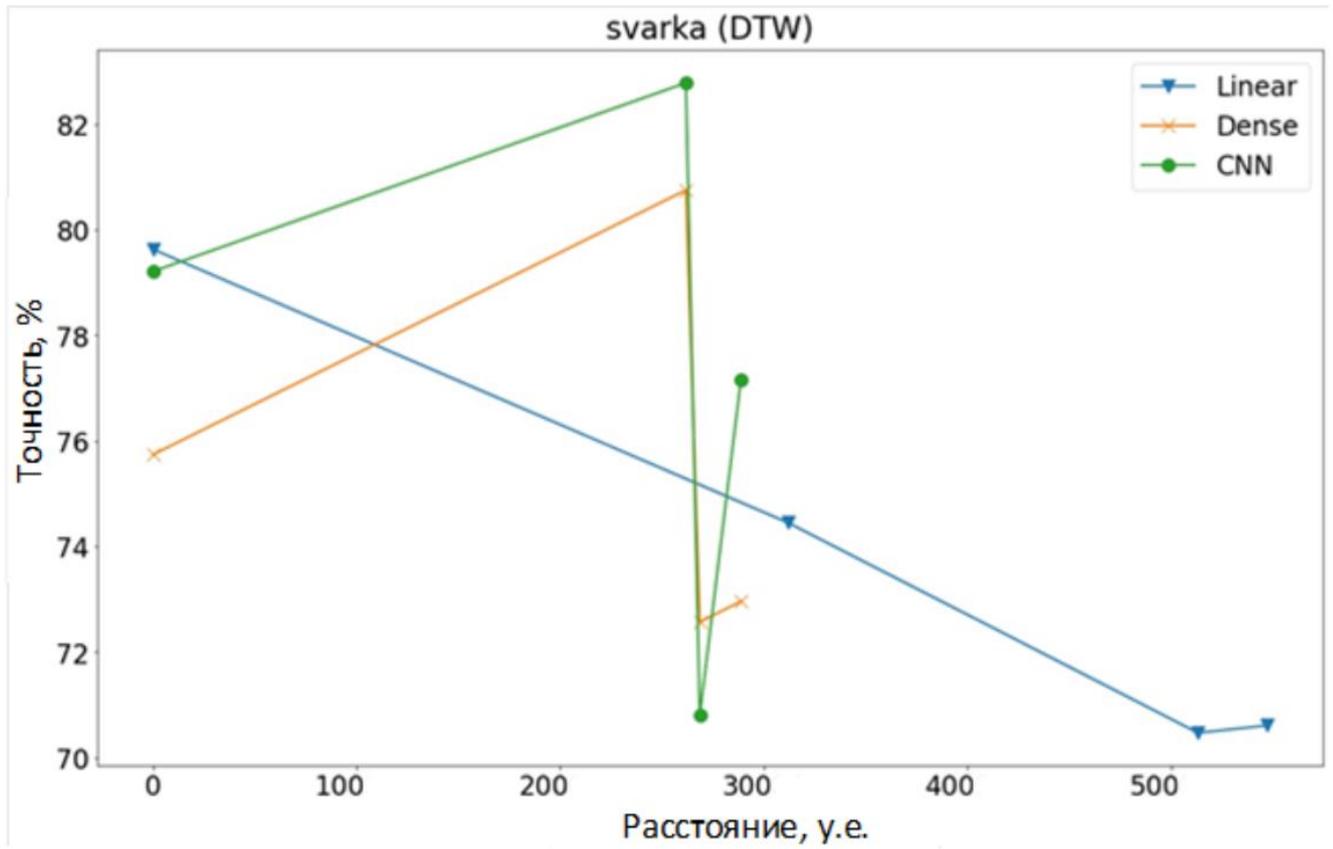


Рисунок 79 – Зависимость точности тестирования моделей от сходства датасетов (данные для обучения модели взяты в работа «svarka»)

Приложение 11 – Алгоритм детектирования аномальных значений (в нотации EPC)

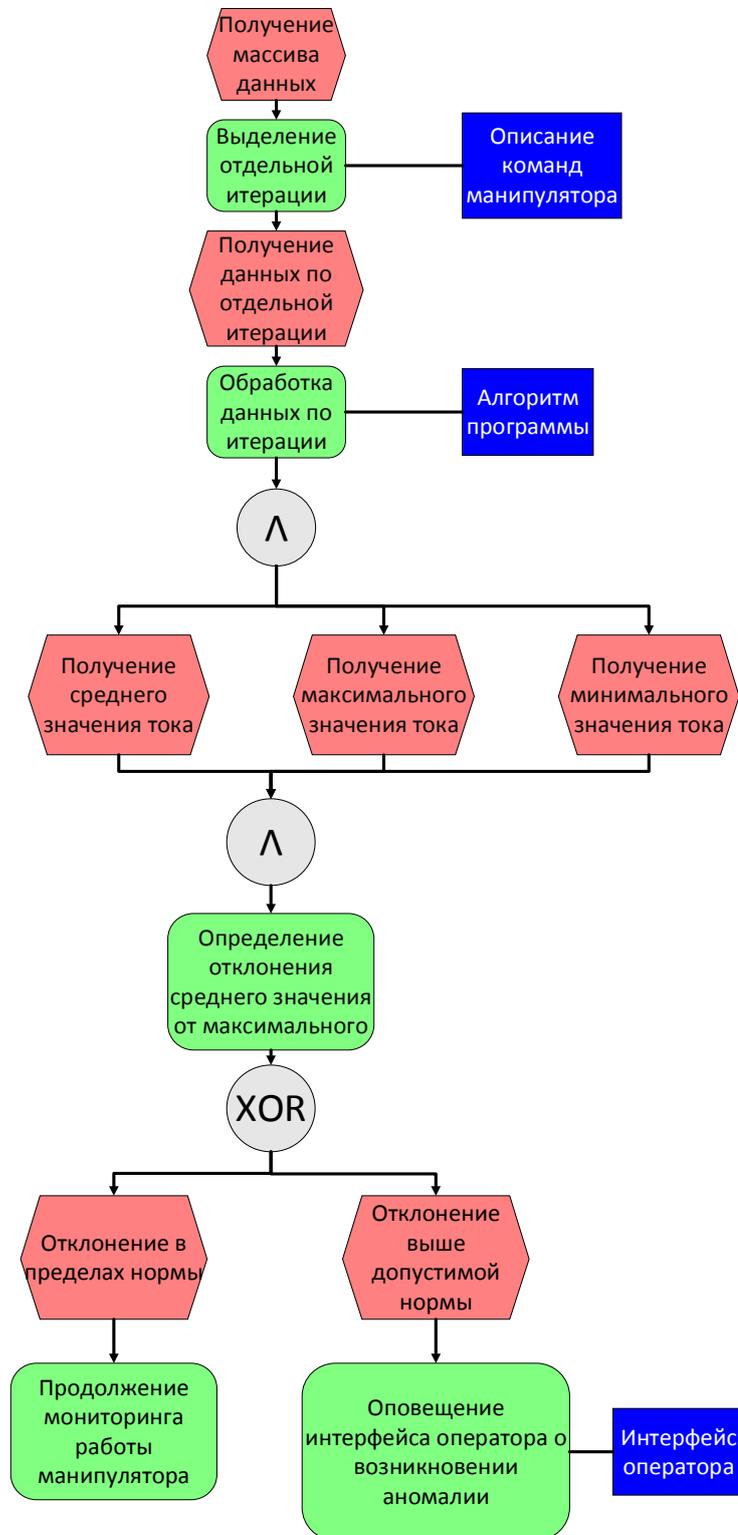


Рисунок 80 – Алгоритм детектирования аномальных значений в работе промышленного манипулятора

Приложение 12 – Акты внедрения результатов диссертационной работы**ООО "Инспайр-технологии"**

ИНН: 7017461845
КПП: 701701001
ОГРН: 1197031059444
ОКПО: 40595188

Расчетный счет: 40702810902500108402
Банк: ООО БАНКА ТОЧКА
БИК: 044525104
Корр. счет: 30101810745374525104

Юридический адрес: 634021, Томская обл, г.о. город Томск, г Томск, ул Шевченко, д. 49Б, стр. 3, офис 20

09.07.2024 г.

Акт внедрения результатов диссертационной работы

Следующие результаты диссертационной работы Гончарова А.С. были использованы при выполнении работ по гранту СТАРТ № 3173ГС1/48576 «Разработка и испытания прототипа программного обеспечения для интеллектуальной диагностики неисправности приводов промышленных манипуляторов» (рег. № НИОКТР АААА-А19-119083090023-2):

- Алгоритмическое и программное обеспечение диагностики аномальных значений приводов промышленных манипуляторов, позволяющее сформировать диапазон допустимых значений при работе, обеспечивающий гибкость процесса мониторинга состояния оборудования по сравнению со штатными функциональными возможностями «KUKA connect».
- Программное обеспечение управления наборами данных о состоянии промышленных роботов и моделями машинного обучения, обеспечивающее визуализацию результатов анализа и обратную связь приводов промышленных манипуляторов.

Директор ООО «Инспайр-технологии»



М.А. Мурин



Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Акт внедрения результатов диссертационной работы

Следующие результаты диссертационной работы Гончарова А.С. были использованы при выполнении работ по гранту РФФИ № 20-37-90113 «Разработка методики адаптивного выбора модели машинного обучения для анализа данных промышленного манипулятора» (руководитель: профессор ОИТ ТПУ Джаякоди Д.Н.К.; рег. № НИОКТР АААА-А20-120082790003-6):

- Алгоритмическое и программное обеспечение рекомендации и оценки эффективности модели машинного обучения для прогнозирования параметров работы промышленного робота, обеспечивающее сокращение временных ресурсов на использование модели анализа данных до 18%, при пороговом значении потери точности до 12%.
- Алгоритмическое и программное обеспечение вычисления меры сходства наборов исторических данных о работе промышленного робота, обеспечивающее подбор модели анализа данных с потерей точности не более 12%.

Проректор по науке
и стратегическим проектам
ФГАОУ ВО НИ ТПУ



А.С. Гоголев



Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)



«УТВЕРЖДАЮ»
Директор ИШИТР, к.т.н.,
А.С. Фадеев
« » 2024 г.

использования результатов диссертационной работы
Гончарова А.С. на тему «Автоматизация управления предиктивным техническим
обслуживанием и ремонтом промышленных роботов» в учебном процессе
Национального исследовательского Томского политехнического университета

Подтверждаем, что результаты диссертационных исследований Гончарова А.С., ассистента отделения информационных технологий и робототехники Инженерной школы информационных технологий и робототехники используются в учебном процессе подготовки студентов бакалавриата отделения информационных технологий и робототехники.

В процессе выполнения диссертационной работы создано алгоритмическое и программное обеспечение рекомендации и оценки эффективности модели машинного обучения для прогнозирования параметров работы промышленного робота, обеспечивающее сокращение временных ресурсов на конструирование модели анализа данных. Материалы использованы в учебном процессе Томского политехнического университета в отделении информационных технологий в качестве:

- Учебно-исследовательской работы студентов бакалавриата по направлению 09.03.04 «Программная инженерия» на тему «Исследование эффективности методов оценки подобию данных и разработка алгоритма подбора обученных моделей».
- Выпускной квалификационной работы студента бакалавриата по направлению 09.03.04 «Программная инженерия» на тему «Анализ данных для прогнозирования отказов и работоспособности промышленного оборудования».

Полученные студентами результаты отражены в отчетах по дисциплине «Учебно-исследовательская работа студента» и опубликованных научных статьях.

Руководитель отделения
информационных технологий на
правах кафедры ИШИТР, к.т.н., доцент

В.С. Шерстнев

Доцент ОИТ ИШИТР, к.т.н.

А.О. Савельев