

На правах рукописи



МИНГАЛИЕВ ЗАИД ЗУЛЬФАТОВИЧ

**МЕТОДЫ И МОДЕЛИ АДАПТИВНОГО ПРОГНОЗИРОВАНИЯ
НЕОДНОРОДНЫХ ВРЕМЕННЫХ РЯДОВ В УСЛОВИЯХ
ОГРАНИЧЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ**

1.2.2. Математическое моделирование, численные
методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Казань – 2026

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Казанский национальный исследовательский технический университет им. А. Н. Туполева-КАИ»

Научный руководитель: доктор технических наук, доцент
Новикова Светлана Владимировна

Официальные оппоненты: **Большаков Александр Афанасьевич**
доктор технических наук, профессор, федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого», профессор Высшей школы технологий искусственного интеллекта;

Орлова Юлия Александровна

доктор технических наук, доцент, федеральное государственное бюджетное образовательное учреждение высшего образования «Волгоградский государственный технический университет», заведующая кафедрой программного обеспечения автоматизированных систем.

Ведущая организация: федеральное государственное бюджетное образовательное учреждение высшего образования «Тверской государственный университет», г. Тверь.

Защита состоится «13» ноября 2026 года в 12:00 на заседании диссертационного совета 24.2.312.08, созданного на базе ФГБОУ ВО «Казанский национальный исследовательский технологический университет» (420015, г. Казань, ул. К. Маркса, д. 68, зал заседаний Ученого совета, А-330).

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Казанский национальный исследовательский технологический университет» и на сайте <https://www.kstu.ru/servlet/contentblob?id=593470>.

Автореферат диссертации разослан «___» _____ 2026 г.

Учёный секретарь
диссертационного совета 24.2.312.08,
доктор экономических наук, доцент



Кудрявцева
Светлана
Сергеевна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Задача прогнозирования временных рядов является одной из ключевых проблем математического моделирования, анализа данных и интеллектуальных систем обработки информации. Временные ряды возникают в широком спектре прикладных областей, включая технические системы, экономику, энергетику, биомедицину, экологический мониторинг и управление сложными динамическими объектами. Для таких данных характерны нестационарность, высокая вариабельность, наличие шумов, пропусков и нелинейных зависимостей, что существенно осложняет построение устойчивых и точных прогнозных моделей.

Современные цифровые системы и киберфизические комплексы предъявляют повышенные требования к методам прогнозирования временных рядов: необходимы высокая точность, адаптивность к изменяющимся условиям, устойчивость к шумам и возможность работы в режиме реального времени при ограниченных вычислительных ресурсах. Эти требования особенно актуальны для интеллектуальных систем поддержки принятия решений и встроенных вычислительных платформ.

В последние годы значительное развитие получили методы прогнозирования временных рядов на основе машинного обучения и глубоких нейросетевых моделей, в частности рекуррентных нейронных сетей и сетей с долговременной краткосрочной памятью (LSTM). Такие модели способны учитывать сложные временные зависимости и демонстрируют высокую точность по сравнению с классическими статистическими методами. Однако универсальные нейросетевые модели часто оказываются недостаточно эффективными при работе с разнородными и структурно различающимися временными рядами, а также требуют значительных вычислительных затрат.

Одним из перспективных направлений повышения качества прогнозирования является предварительная структуризация временных рядов, в частности их кластеризация по динамическим и статистическим характеристикам, а также снижение размерности данных за счет адаптивной аппроксимации. Комбинация методов кластеризации, предварительной обработки и специализированных прогнозных моделей позволяет учитывать неоднородность данных, повышать устойчивость прогнозов и снижать вычислительную сложность алгоритмов.

В качестве прикладных областей апробации разработанных методов в диссертационной работе рассматриваются физиологические временные ряды, в том числе данные непрерывного мониторинга уровня глюкозы у пациентов с сахарным диабетом, а также временные ряды оптимального управления беспилотными системами в условиях внешних помех. Данные предметные области являются актуальными, но не ограничивающими: предложенные модели и алгоритмы ориентированы на широкий класс задач прогнозирования временных рядов и могут быть применены в различных областях науки и техники.

Таким образом, актуальной научной задачей является разработка методов и моделей адаптивного прогнозирования временных рядов, основанных на их предварительной кластеризации и использовании нейросетевых архитектур, обеспечивающих повышение точности прогнозов и снижение вычислительных затрат.

Степень разработанности темы. Современные методы прогнозирования временных рядов условно делятся на три основные категории: классические статистические модели, методы машинного обучения и глубокие нейросетевые алгоритмы.

Классические подходы, такие как модели ARIMA, экспоненциальное сглаживание и фильтр Калмана, нашли широкое применение в анализе временных рядов. Работы

Бокса и Дженкинса, а также Калмана заложили основу для статистического моделирования линейных временных зависимостей. Однако, как отмечают Хайндман и Афанасопулос, данные методы плохо справляются с нелинейными и высоко вариативными данными.

Методы машинного обучения, включая градиентный бустинг, случайные леса и метод опорных векторов, продемонстрировали более высокую гибкость в задачах прогнозирования. В работах Бреймана, Фридмана и Кушнарева показано, что такие алгоритмы способны выявлять сложные закономерности, особенно при наличии большого количества признаков. Однако они требуют значительных объёмов размеченных данных и тщательной предварительной обработки, что ограничивает их применение в реальных условиях.

Глубокие нейросетевые модели, такие как LSTM (Хохрайтер и Шмидхубер), сверточные нейросети (Лекун) и скрытые марковские модели (Рабинер), на сегодняшний день демонстрируют наивысшую точность при прогнозировании временных рядов. Особенно эффективными они оказались в задачах анализа гликемии, о чём свидетельствуют исследования Чжу и Перес-Гандия, где LSTM-сети использовались для предсказания уровня глюкозы с учётом предыдущих измерений, данных о питании и физической активности. Однако высокая вычислительная сложность таких моделей и их потребность в значительных вычислительных ресурсах затрудняют применение в условиях реального времени, особенно в мобильных, носимых или легких бортовых устройствах.

Несмотря на значительный прогресс, в задачах прогнозирования временных рядов сохраняется проблема адаптации моделей к индивидуальным и контекстно-зависимым особенностям данных. Универсальные прогнозные модели не всегда обеспечивают требуемую точность при работе с разнородными и нестационарными временными рядами. Это характерно как для прогнозирования физиологических показателей (в частности, уровня глюкозы у пациентов с сахарным диабетом), так и для задач оптимального управления динамическими объектами, включая беспилотные летательные аппараты, где динамика данных определяется режимами функционирования и внешними возмущениями. В связи с этим актуальной является разработка адаптивных методов прогнозирования, учитывающих характерные паттерны конкретных временных рядов.

Ещё одна важная задача - работа с зашумлёнными и разреженными данными. Данные, получаемые с помощью внешних сенсоров, могут содержать пропуски, артефакты и быть получены с различной частотой. Эти вопросы поднимаются в исследованиях Бунеску и Дзеккин, где рассматриваются методы предварительной обработки данных, полученных с внешних сенсорных устройств.

В работе предложен комплексный подход к прогнозированию временных рядов, направленный на повышение точности и адаптивности предсказаний. Использование предварительной кластеризации временных рядов позволяет учитывать структурные особенности данных, формировать более однородные группы наблюдений и снижать вычислительную нагрузку за счёт обучения специализированных прогнозных моделей для каждого кластера.

Цель и задачи исследования.

Целью работы является повышение точности и адаптация прогнозов к индивидуальной динамике неоднородных временных рядов.

Для достижения поставленной цели в работе решаются следующие задачи:

- анализ и обзор современных методов прогнозирования временных рядов, включая статистические, машинно-обучающие и нейросетевые подходы;

- разработка метода адаптивной кусочно-линейной аппроксимации временных рядов с динамическим выбором шага;
- разработка метода кластеризации временных рядов на основе вычисления статистических и динамических характеристик, отражающих их динамическую структуру;
- разработка модели адаптивного прогнозирования временных рядов на основе LSTM в комбинации с методом кластеризации временных рядов и методов кусочно-линейной аппроксимации;
- реализация программного комплекса, обеспечивающего полный цикл обработки временных рядов: предобработку, кластеризацию и прогнозирование;
- проведение вычислительных экспериментов для оценки точности, устойчивости и вычислительной эффективности разработанных методов;
- апробация предложенных решений на реальных временных рядах в различных предметных областях.

Актуальной **научной задачей**, решаемой в диссертации, является разработка методов и моделей адаптивного прогнозирования неоднородных временных рядов, а также программного обеспечения для их практической реализации в условиях ограниченных вычислительных ресурсов, на основе глубокого обучения.

Объект исследования: временные ряды, характеризующие динамику сложных процессов.

Предмет исследования: методы и модели адаптивного прогнозирования временных рядов на основе глубокого обучения.

Новизна полученных результатов заключается в следующем:

Научная новизна работы заключается в интеграции методов кластеризации временных рядов с глубокими нейросетевыми моделями, оптимизированными для работы в условиях ограниченных вычислительных ресурсов, а именно:

1. Разработан метод адаптивной кластеризации временных рядов, базирующийся на анализе их статистических и динамических характеристик, отличающийся возможностью выявления однородных групп временных рядов с учетом индивидуальных и контекстно-зависимых особенностей их динамики, что позволяет повысить точность последующего прогнозирования (п. 2 паспорта специальности 1.2.2).

2. Разработан метод адаптивной кусочно-линейной аппроксимации временных рядов, базирующийся на сегментации данных с переменным шагом по заданному критерию точности, отличающийся снижением размерности данных при сохранении их структурных свойств, что позволяет уменьшить вычислительные затраты на последующих этапах обработки и прогнозирования (п. 2 паспорта специальности 1.2.2).

3. Разработана гибридная нейросетевая модель прогнозирования временных рядов, базирующаяся на интеграции LSTM-архитектур, методов кластеризации и адаптивной аппроксимации, отличающаяся использованием предварительной структуризации данных для формирования скрытых представлений, что позволяет повысить устойчивость прогнозов к шумам и пропускам, а также снизить вычислительную сложность модели (п. 8 паспорта специальности 1.2.2).

4. Разработан программный комплекс для прогнозирования временных рядов, базирующийся на интеграции методов кластеризации, адаптивной кусочно-линейной аппроксимации и гибридных нейросетевых моделей, отличающийся модульной архитектурой и возможностью настройки под различные классы временных данных, что позволяет реализовать адаптивный подход к прогнозированию в системах реального времени (п. 3 паспорта специальности 1.2.2).

5. Проведены численные эксперименты по решению задач прогнозирования неоднородных временных рядов, базирующиеся на использовании синтетических и реальных наборов данных с последующим статистическим анализом результатов, отличающиеся комплексной оценкой точности, устойчивости и вычислительной эффективности моделей, что позволяет обоснованно оценить применимость и эффективность разработанных методов и алгоритмов для решения поставленных задач (п. 9 паспорта специальности 1.2.2).

В отличие от существующих универсальных моделей прогнозирования, предложенный подход обеспечивает учет структурной неоднородности временных рядов, снижение вычислительных затрат и повышение устойчивости прогнозов в условиях шумов и пропусков данных.

Теоретическая значимость работы состоит в развитии методов анализа и прогнозирования временных рядов с учетом их структурных и вероятностных свойств, а также в обосновании эффективности предварительной кластеризации для повышения качества нейросетевого прогнозирования.

Практическая значимость заключается в возможности применения разработанных методов в интеллектуальных системах прогнозирования и поддержки принятия решений в различных предметных областях. В качестве прикладного примера показана эффективность предложенного подхода при прогнозировании физиологических временных рядов, включая данные мониторинга уровня глюкозы, а также рядов оптимального управления беспилотными аппаратами.

Методология и методы исследования. Для решения поставленных задач в диссертационной работе использовались предиктивные методы аппроксимации и регрессии, математической статистики, математического и нейросетевого моделирования, кластерного анализа, а также методы оптимизации и объектно-ориентированного программирования.

Соответствие диссертации паспорту научной специальности. В диссертации разработаны и реализованы в виде комплекса программ модели и методы предварительной обработки и прогнозирования временных рядов. Для оценки их адекватности и эффективности приведены результаты вычислительных экспериментов.

Таким образом, результаты диссертационного исследования соответствуют следующим пунктам паспорта научной специальности 1.2.2. Математическое моделирование, численные методы и комплексы программ:

п. 2. *Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий:* разработаны, обоснованы и протестированы вычислительные методы адаптивной кластеризации и кусочно-линейной аппроксимации временных рядов;

п. 3. *Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента:* реализация разработанных методов и алгоритмов в виде комплекса программ, на базе которого проведены вычислительные эксперименты по оценке эффективности разработанного алгоритмического обеспечения и возможности его практического использования для решения поставленных задач;

п. 8. *Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента:* разработана гибридная нейросетевая модель прогнозирования временных рядов с применением адаптивных методов предварительной кластеризации и кусочно-линейной аппроксимации;

п. 9. *Постановка и проведение численных экспериментов, статистический анализ их результатов, в том числе с применением современных компьютерных технологий*: проведены численные эксперименты с последующим статистическим анализом их результатов для оценки эффективности разработанных методов и алгоритмов.

Положения, выносимые на защиту:

1. Метод адаптивной кластеризации временных рядов, позволяющий группировать временные ряды различной длины и структуры для повышения точности последующих прогнозов.

2. Метод адаптивной кусочно-линейной аппроксимации временных рядов, позволяющий уменьшить вычислительные затраты при прогнозировании.

3. Гибридная нейросетевая модель прогнозирования временных рядов, снижающая вычислительную сложность и повышающая устойчивость прогнозов к шумам и пропускам.

4. Архитектура оригинального комплекса программ, позволяющая реализовать разработанные модели, методы и алгоритмы в среде C# (Visual Studio Community 2022) и Python (PyCharm 2022).

Степень достоверности и апробация результатов. Достоверность научных результатов подтверждена вычислительными экспериментами и результатами практического использования. Диссертационная работа выполнена при поддержке Министерства науки и высшего образования в рамках исполнения обязательств по НИОКТР 126020516513–4, рег. номер FZSU–2026–0007.

Апробация работы. Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих конференциях: Современные цифровые технологии: проблемы, решения, перспективы: национальная (с международным участием) научно-практическая конференция, Казань, 19–20 мая 2022 года; Цифровые системы и модели: теория и практика проектирования, разработки и использования: Международная научно-практическая конференция, Казань, 10–11 апреля 2025 года; Профессиональные коммуникации в научной среде – фактор обеспечения качества исследований: XV Всероссийская научно-практическая конференция, Альметьевск, 16 апреля 2025 года; XXVI Туполевские чтения (школа молодых ученых): Международная молодёжная научная конференция, Казань, 09–10 ноября 2023 года; Приоритетные направления инновационной деятельности в промышленности: сборник научных статей по итогам девятой международной научной конференции, Казань, 29–30 сентября 2020 года; Кулагинские чтения: техника и технологии производственных процессов: XX Международная научно-практическая конференция, Чита, 30 ноября – 04 декабря 2020 года.

Реализация результатов диссертации:

1) внедрены в ФГБОУ ВО «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ» и используются в учебном процессе при изучении дисциплин для подготовки магистров по направлению «Прикладная математика и информатика» при проведении лекционных занятий и лабораторных работ по дисциплинам «Компьютерное моделирование процессов и систем», «Теория принятия решений» и «Теория игр и исследование операций»;

2) использованы в работе Сектора безопасности дорожного движения и контрольно-надзорной деятельности Министерства Транспорта и дорожного хозяйства Республики Татарстан в системе поддержки принятия решений для моделирования траекторий движения транспортных потоков на дорожных развязках сложных конфигураций на территории Республики Татарстан;

3) использованы при реализации научного направления деятельности Центра разработки и сопровождения информационных систем ГБУ «Безопасность дорожного движения» при разработке научных и практических основ моделирования транспортных потоков;

4) использованы в работе отделения эндокринологии в системе поддержки принятия решений для прогнозирования уровня гликемии у больных инсулинозависимым диабетом, получающих помповую инсулинотерапию;

5) использованы в работе Научно-производственного центра беспилотных авиационных систем (НПЦ БАС) Татарстана для имитации сложных траекторий движения и моделирования управления беспилотных летательных аппаратов.

Публикации. Материалы диссертации опубликованы в 12 научных работах общим объемом 4,38 п.л. (в т.ч. лично автором – 3,1 п.л.), из них 5 статей в журналах из перечня рецензируемых научных изданий, рекомендованных ВАК Министерства науки и высшего образования Российской Федерации – «Современные информационные технологии и ИТ-образование», «Вестник Тверского государственного университета», «Вестник Технологического университета»; 1 статья в издании, которое входит в международные реферативные базы данных и систем цитирования Scopus. Зарегистрированы 4 свидетельства о государственной регистрации программ для ЭВМ (модули «Адаптивный модуль аппроксимации», «Программа анализа и прогнозирования смешанных временных рядов», «Программа кластеризации и адаптивного прогнозирования временных рядов на основе LSTM», «Программа синтаксического анализа больших данных биологического мониторинга»).

Структура и объем диссертации.

Диссертация изложена на 163 страницах машинописного текста, содержит 38 рисунков, 33 таблицы, состоит из введения, четырех глав, заключения, списка использованной литературы из 118 наименований и приложений.

Сведения о личном вкладе автора. Постановка научной задачи, содержание диссертации и все представленные в ней результаты получены лично автором. Подготовка к публикации некоторых результатов проводилась совместно с соавторами, причем вклад диссертанта определяющий.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении содержится обоснование актуальности задачи исследования, приводятся основные научные положения и результаты.

В первой главе проведен обзор существующих моделей и методов кластеризации и прогнозирования временных рядов. Рассмотрены как классические, так и современные подходы к анализу временных рядов с учетом их специфики: высокой изменчивости, пропущенных значений и малой информативности отдельных измерений.

Проанализированы методы кластеризации временных рядов. Приведен обзор подходов на основе измерения расстояний, включая метрики DTW, Euclidean и другие, а также методов вероятностного моделирования, таких как скрытые марковские модели и байесовские подходы. Особое внимание уделено современным методам глубокого обучения, включая автоэнкодеры и рекуррентные нейронные сети, позволяющие выявлять скрытые структуры и закономерности во временных рядах.

Рассмотрены методы прогнозирования временных рядов. Проанализированы классические статистические модели (AR, ARIMA, экспоненциальное сглаживание), методы машинного обучения (регрессии, случайные леса, градиентный бустинг) и современные глубокие нейросетевые модели (LSTM, GRU, трансформеры). Подчеркнута их

способность моделировать сложные нелинейные зависимости, учитывать динамику гликемии и адаптироваться к особенностям пациентов.

Проведен анализ преимуществ и ограничений различных методов, выявлены проблемы в применении существующих подходов для прогнозирования неоднородных временных рядов. Обоснована необходимость разработки специализированных алгоритмов и программного комплекса, способных адаптивно кластеризовать и прогнозировать неоднородные временные ряды с учетом индивидуальных и контекстно-зависимых особенностей данных. Сформулированы задачи исследования.

Во второй главе представлены разработанные методы кластеризации и прогнозирования временных рядов. Предложен метод адаптивной кусочно-линейной аппроксимации с переменным шагом, обеспечивающий снижение вычислительных затрат при сохранении заданной точности. Разработан метод кластеризации временных рядов на основе вычисления статистических и динамических характеристик с использованием алгоритма DBSCAN, позволяющий формировать однородные группы данных. Для прогнозирования предложена адаптивная LSTM-модель, обучаемая отдельно для каждого кластера, что обеспечивает повышение точности предсказаний. Эффективность предложенного подхода подтверждена сравнительным анализом, результаты которого представлены в Главе 4.

I. Метод адаптивной кусочно-линейной аппроксимации временных рядов.

Данный метод позволяет рассчитывать переменный шаг для кусочно-линейной аппроксимации эмпирической нелинейной функции с заданной точностью. Рассмотрим эмпирическую нелинейную одномерную функцию $y = f(x)$, заданную таблично на интервале $[a_1; a_n]$ в виде пары векторов-столбцов координат входных значений $X^n = (x_1, x_2 \dots x_n)^T$ и соответствующих координат выходных значений $Y^n = (y_1, y_2 \dots y_n)^T$. Здесь $x_1 = a_1$, $x_n = a_n$, n – количество опорных точек-узлов таблично заданной функции.

Пары «вход-выход» образуют матрицу T размерности $(n \times 2)$ следующего вида:

$$T = \begin{pmatrix} x_1 & y_1 \\ \dots & \dots \\ x_n & y_n \end{pmatrix} = [X^n | Y^n] \quad (1)$$

Введем понятие *сегмента*.

Сегмент – это подматрица матрицы T , включающая в себя последовательные строки матрицы T с некоего номера u до номера v . Сегменты не пересекаются между собой. В совокупности сегменты образуют матрицу T .

Обозначим сегмент следующим образом:

$$S^m(u, v) = \begin{pmatrix} x_u & y_u \\ \dots & \dots \\ x_v & y_v \end{pmatrix} \quad (2)$$

$$S^m \cap S^l = \emptyset \quad (3)$$

$$\bigcup_{m=1}^Q S^m = T \quad (4)$$

Здесь: Q – количество сегментов, m – порядковый номер сегмента, u – индекс начального элемента сегмента, v – индекс конечного элемента сегмента.

Необходимо декомпозировать исходную матрицу T на сегменты таким образом, чтобы при замене участка нелинейной функции внутри сегмента линейным отрезком

результатирующая кусочно-линейная функция аппроксимировала исходную нелинейную функцию с заданной точностью. При этом в качестве размера шага аппроксимации h^m принимается длина соответствующего сегмента $h^m = d[S^m(u, v)]$.

Для решения поставленной задачи был разработан численный метод определения на матрице Т сегментов максимально возможной длины при условии, что внутри сегмента линейный многочлен аппроксимирует исходную функцию с заданной точностью ε . В качестве критерия точности используется коэффициент детерминации R^2 , вычисляемый для линейной регрессионной функции внутри сегмента. В ходе исполнения вычислений производится два цикла вычислений: внешний цикл отвечает за сокращение исходного интервала поиска $[a_1, a_n]$ по мере того, как будут выделяться последовательные сегменты аппроксимации; внутренний цикл отвечает за выделение границ очередного сегмента внутри интервала поиска.

Метод адаптивной кусочно-линейной аппроксимации временных рядов состоит из следующих шагов:

Шаг 0. Установить начальные значения: номер определяемого сегмента $m=1$, индекс первого элемента определяемого сегмента $u=1$. Задать параметр точности ε .

Начало внешнего цикла

Шаг 1. $v=n$

Начало внутреннего цикла

Шаг 2. Для сегмента $S^m(u, v)$ построить линейную регрессионную функцию $\hat{y} = A + Bx$ с коэффициентами, определяемыми согласно методу наименьших квадратов:

$$A = \frac{\sum_{i=u}^v y_i \sum_{i=u}^v x_i^2 - \sum_{i=u}^v x_i \sum_{i=u}^v x_i y_i}{(v - u + 1) \sum_{i=u}^v x_i^2 - (\sum_{i=u}^v x_i)^2} \quad (5)$$

$$B = \frac{(v - u + 1) \sum_{i=u}^v x_i y_i - \sum_{i=u}^v x_i \sum_{i=u}^v y_i}{(v - u + 1) \sum_{i=u}^v x_i^2 - (\sum_{i=u}^v x_i)^2} \quad (6)$$

Шаг 3. Вычислить значения регрессионной функции в узлах сетки сегмента $S^m(u, v)$:

$$\hat{y}_i = A + Bx_i \quad (7)$$

$$i = \overline{u, v} \quad (8)$$

Шаг 4. Вычисляется коэффициент детерминации для сегмента $S^m(u, v)$:

$$R^2 = 1 - \frac{\sum_{i=u}^v (y_i - \hat{y}_i)^2}{\sum_{i=u}^v (y_i - \bar{y})^2} \quad (9)$$

$$\bar{y} = \frac{\sum_{i=u}^v y_i}{(v - u + 1)} \quad (10)$$

Шаг 5. Если $R^2 < \varepsilon$ то $v = v - 1$. Возврат к шагу 2. Иначе переход к шагу 6.

Конец внутреннего цикла.

Шаг 6. Если $v < n$, то $m = m + 1$, $u = v + 1$, возврат к Шагу 1. Иначе – $Q = m$, завершить формирование сегментов.

Конец внешнего цикла.

II. Метод адаптивной кластеризации временных рядов.

Метод позволяет кластеризовать временные ряды на основе вычисления статистических и динамических характеристик ряда, и направлен на улучшение точности группировки данных за счет выделения ключевых параметров временных серий. В отличие от традиционных подходов, использующих расстояние или преобразование данных в векторы признаков, данный метод учитывает временную структуру последовательностей.

Пусть сегмент временного ряда задаётся последовательностью:

$$y = \{y_1, y_2, \dots, y_L\}, \quad (11)$$

где y_t - значение показателя в момент времени t .

Метод адаптивной кластеризации временных рядов с последующим построением прогнозирующих моделей состоит из следующих шагов:

Шаг 1. Ряд делится на сегменты длиной L точек.

Шаг 2. Заполнение пропущенных значений. Используется линейная интерполяция между соседними наблюдаемыми точками.

Шаг 3. Сглаживание шума. Применяется скользящее среднее с окном размера w :

$$\hat{y}_t = \frac{1}{w} \sum_{i=t-w+1}^t y_i, \quad (12)$$

где $\hat{y}_t$ — сглаженное значение в точке t . Это уменьшает влияние краткосрочных артефактов, не искажая долгосрочные тренды.

Шаг 4. Для каждого сегмента вычисляется вектор временных характеристик: минимум, максимум (13), среднее значение (14), медианное значение (15), дисперсия (16), коэффициент асимметрии (17), коэффициент эксцесса (18), автокорреляция (19), площадь под графиком (20), наклон тренда (21).

$$y_{min} = \min_{1 \leq t \leq L} y_t, y_{max} = \max_{1 \leq t \leq L} y_t \quad (13)$$

$$\mu = \frac{1}{L} \sum_{t=1}^L y_t \quad (14)$$

$$med(y) = \begin{cases} y_{(k)}, L = 2k - 1 \\ \frac{y_{(k)} + y_{(k+1)}}{2}, L = 2k \end{cases} \quad (15)$$

$$\sigma^2 = \frac{1}{L} \sum_{t=1}^L (y_t - \mu)^2 \quad (16)$$

$$\gamma_1 = \frac{1}{L} \sum_{t=1}^L \left(\frac{y_t - \mu}{\sigma} \right)^3 \quad (17)$$

$$\gamma_2 = \frac{1}{L} \sum_{t=1}^L \left(\frac{y_t - \mu}{\sigma} \right)^4 - 3 \quad (18)$$

$$r_1 = \frac{\sum_{t=2}^L (y_t - \mu)(y_{t-1} - \mu)}{\sum_{t=1}^L (y_t - \mu)^2} \quad (19)$$

$$S = \sum_{t=1}^L y_t \Delta t \quad (20)$$

$$y_t = at + b + \varepsilon_t \quad (21)$$

Шаг 5. Применяем масштабирование для приведения всех признаков к единому масштабу.

Шаг 6. Выполняем кластеризацию полученного массива векторов алгоритмом DBSCAN.

Шаг 7. Для каждого полученного кластера отдельно формируем обучающую выборку из временных рядов данного кластера. Выполняем обучение LSTM-сети с заданными гиперпараметрами и валидацию модели на отложенной выборке.

Шаг 8. При прогнозировании для нового временного ряда сначала определяется его кластер с помощью модели DBSCAN, а затем для соответствующего кластера используется обученная LSTM-модель для предсказания значений временного ряда.

Таким образом, предложенная методология решает проблему кластеризации и прогнозирования временных рядов, учитывая индивидуальные особенности данных и позволяя более точно моделировать динамику для различных типов временных рядов.

III. Гибридная нейросетевая модель прогнозирования временных рядов.

Для каждого кластера, полученного на предыдущем этапе кластеризации, обучается отдельная модель LSTM, что позволяет более точно учитывать особенности каждого типа данных. Такой подход позволяет адаптировать нейросети под специфические зависимости временных рядов в рамках каждого кластера, что повышает точность прогнозирования.

Когда требуется сделать прогноз для нового временного ряда, процесс начинается с его классификации. С помощью модели кластеризации временных рядов на основе временных характеристик определяется, к какому кластеру принадлежит новый ряд. После того как кластер определен, для соответствующего кластера используется соответствующая модель LSTM, обученная на данных этого кластера. Это позволяет эффективно использовать особенности каждого кластера, улучшая точность предсказания.

Такой подход решает задачу прогнозирования, разделяя временные ряды на более однородные группы, что позволяет каждой модели LSTM работать с более похожими данными и, соответственно, делать более точные предсказания.

Архитектура разработанной гибридной модели LSTM для прогнозирования временного ряда представлена на рисунке 1.

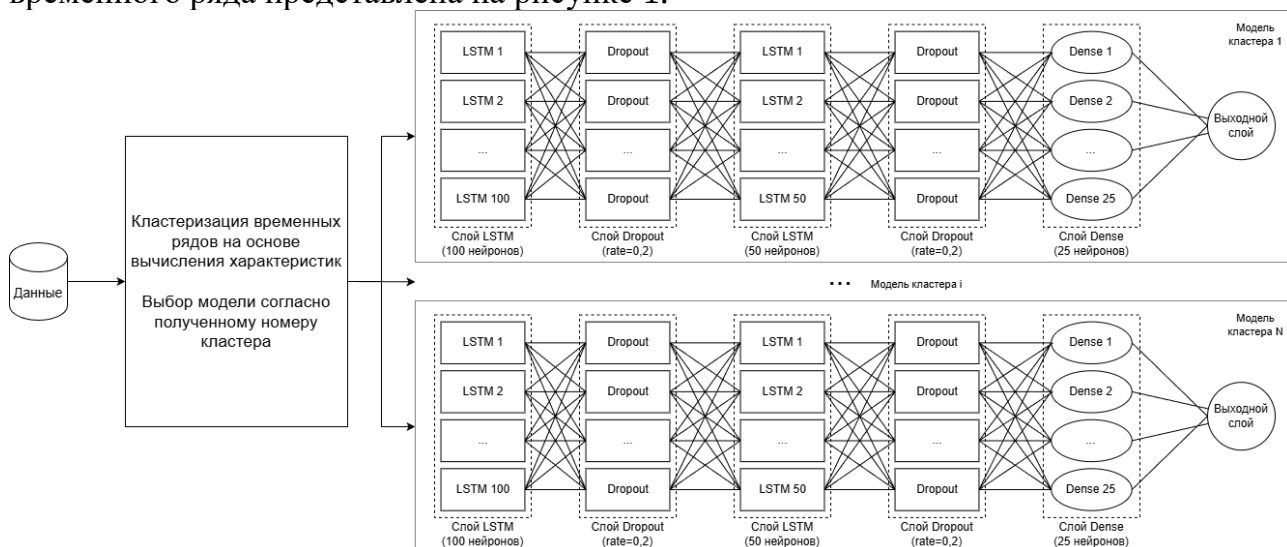


Рисунок 1 – Архитектура LSTM-модели

В третьей главе представлены разработанные программные комплексы, реализующие методы и модели кластеризации и прогнозирования временных рядов. Для апробирования метода расчета переменного шага для кусочно-линейной аппроксимации создан адаптивный модуль аппроксимации, реализующий как жадный, так и адаптивный режимы разбиения функций на сегменты с сохранением заданной точности. Программный модуль позволяет генерировать электронные документы с метриками сегментов. Модуль разработан на платформе .NET 7 с использованием C# и инструментов Visual Studio 2022, и зарегистрирован в качестве программы для ЭВМ.

Создан программный комплекс для прогнозирования временных рядов с предварительной кластеризацией. Комплекс включает блоки предобработки данных, обучения моделей LSTM для каждого кластера и прогнозирования новых временных рядов с использованием соответствующей модели. Для кластеризации применяется алгоритм DBSCAN, а параметры временных рядов учитывают статистические и динамические характеристики. Для предобработки данных использован Python с библиотеками pandas и numpy, а интеграция с серверной частью выполнена на ASP.NET Core и C#.

Разработанный программный комплекс обеспечивает модульность, высокую производительность и возможность адаптации под различные задачи и типы временных данных. Архитектура программного комплекса представлена на рисунке 2.

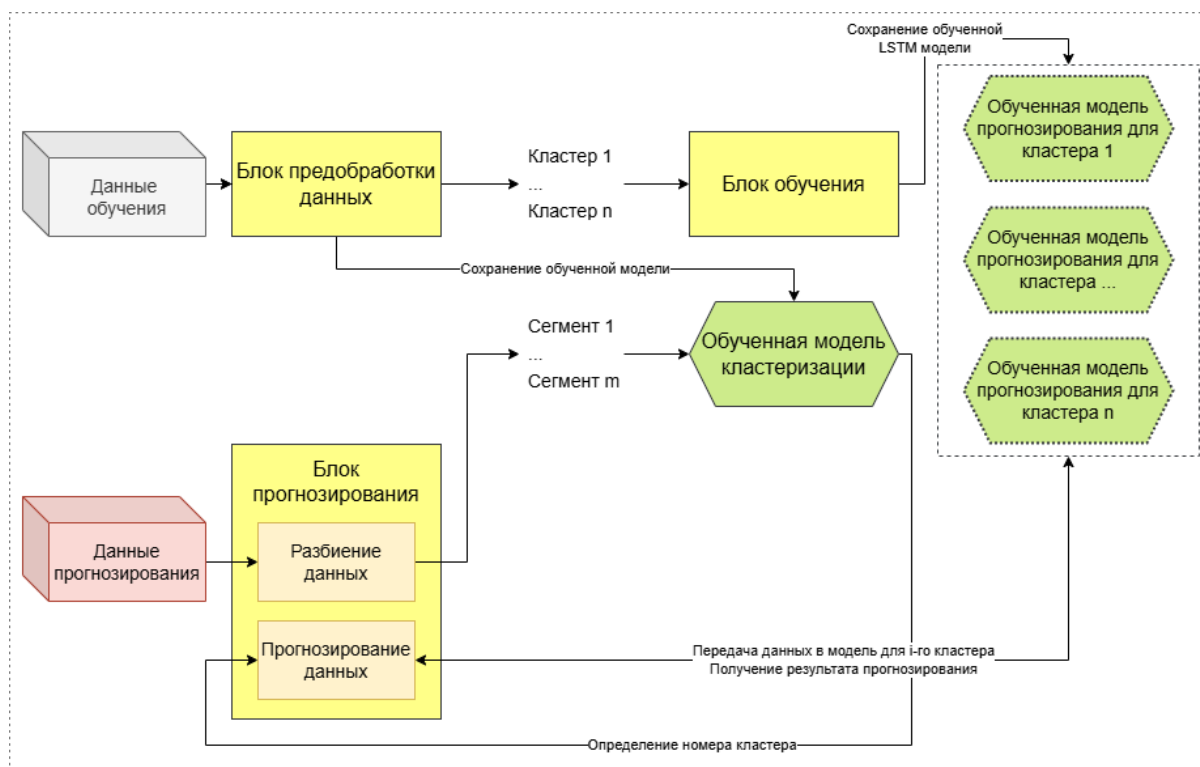


Рисунок 2 – Верхнеуровневая архитектура программного комплекса

На все разработанные модули программного комплекса получены свидетельства о государственной регистрации программ для ЭВМ.

В четвертой главе описаны проведенные вычислительные эксперименты для тестирования применимости и точности разработанных методов и алгоритмов в различных предметных областях.

В разделе 4.1 описываются вычислительные эксперименты по применению метода адаптивной кусочно-линейной аппроксимации временных рядов. В первой части раздела метод применен для аппроксимации законов управления гибридным

летательно-подводным аппаратом типа окто-квадрокоптер. Рассматривались траектории управления по углам тангажа, крена и рысканья при переходе аппарата между средами «вода–воздух». Исходные траектории были оцифрованы с шагом 0,5 с и сглажены методом экспоненциального сглаживания. Затем был применен предложенный метод адаптивной кусочно-линейной аппроксимации временных рядов. Расхождение графиков реальных и рассчитанных кусочно-линейных траекторий по трем каналам управления демонстрирует рисунке 3.

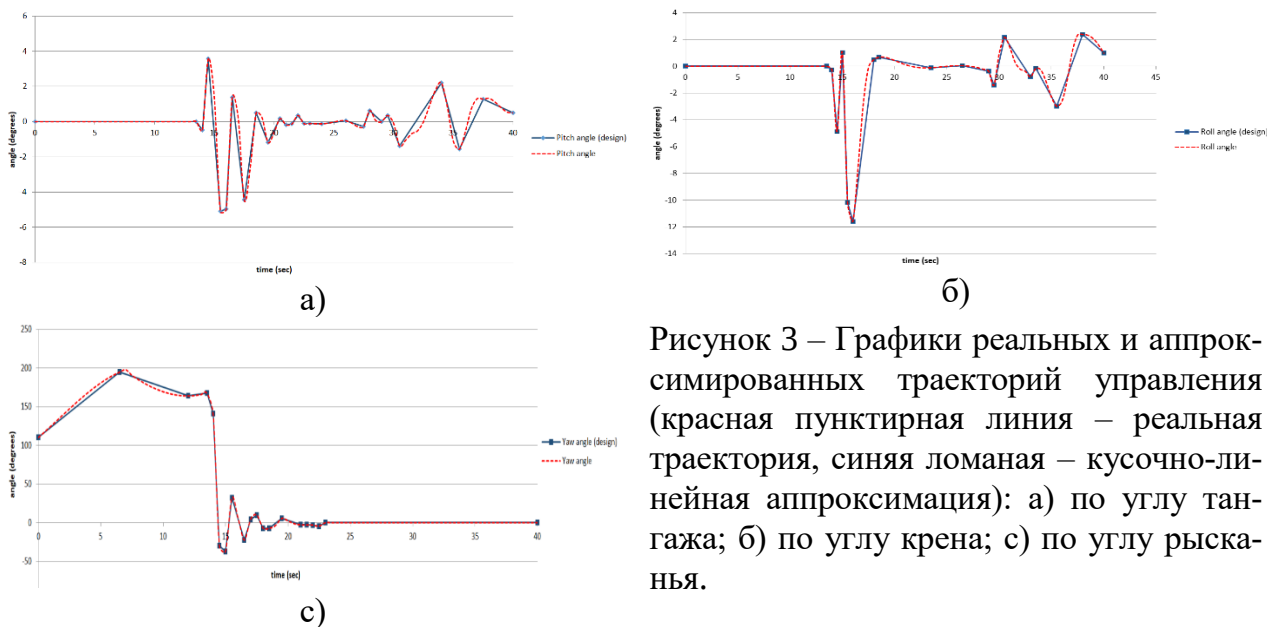


Рисунок 3 – Графики реальных и аппроксимированных траекторий управления (красная пунктирная линия – реальная траектория, синяя ломаная – кусочно-линейная аппроксимация): а) по углу тангажа; б) по углу крена; в) по углу рысканья.

Предложенный метод адаптивной кусочно-линейной аппроксимации временных рядов с переменным шагом сравнивался с базовым подходом с постоянным шагом при одинаковой точности аппроксимации (коэффициент детерминации не ниже $R^2 \geq 0,9$). В качестве критериев сравнения использовались объем требуемой памяти и количество операций, необходимых для реализации управления на бортовом компьютере. Сравнение ресурсных затрат для двух подходов приведено в таблицах 1 и 2.

Таблица 1 – Сравнение объема хранимых параметров управления

Канал управления	Переменный шаг	Постоянный шаг	Сокращение
Тангаж	104	241	2,3 раза
Крен	72	241	3,3 раза
Рысканье	72	241	3,3 раза
Итого	248	723	≈3 раза

Таблица 2 – Сравнение количества операций сравнения и переключения

Канал управления	Переменный шаг	Постоянный шаг	Сокращение
Тангаж	2132	6560	3,1 раза
Крен	1476	6560	4,4 раза
Рысканье	1476	6560	4,4 раза
Итого	5084	19680	≈4 раза

Показано, что при использовании переменного шага число сегментов аппроксимации составило 26 для тангажа и по 18 для крена и рысканья, в то время как метод с постоянным шагом требовал 80 сегментов для каждого канала управления. В результате адаптивный метод обеспечивает сокращение объема хранимых параметров

примерно в 3 раза и уменьшение количества операций управления почти в 4 раза по сравнению с методом с постоянным шагом при сохранении заданной точности аппроксимации.

Затем предложенный метод применен для кусочно-линейной аппроксимации нелинейной функции, описывающей динамику концентрации экзогенного инсулина в плазме крови пациента с сахарным диабетом 1 типа. Рассматривалась кривая активности инсулина ультракороткого действия продолжительностью 6 часов. Адаптивный метод с переменным шагом сравнивался с методом с постоянным шагом. Результаты вычислительных экспериментов показали, что при одинаковой точности аппроксимации предложенный метод адаптивной кусочно-линейной аппроксимации с переменным шагом требует в 3 раза меньше сегментов и более чем в 11 раз меньше вычислительных операций по сравнению с методом с постоянным шагом (таблица 3).

Таблица 3 – Сравнение вычислительных затрат

Показатель	Переменный шаг	Постоянный шаг	Сокращение
Точность аппроксимации	$\geq 0,99$	$\geq 0,99$	одинаковая
Кол-во сегментов	5	15	3 раза
Число операций (всего)	825	9316	11,3 раза
Основной вклад	операции сравнения	арифметические операции	-

В разделе 4.2. описаны эксперименты по применению разработанного метода адаптивной кластеризации и гибридной модели прогнозирования временных рядов. Проведены вычислительные эксперименты по формированию, обучению и тестированию многопоточного нейросетевого контроллера для расчёта управляющих параметров автопилота легкого самолета при развороте. Исходный датасет временных рядов был предварительно кластеризован предложенным методом адаптивной кластеризации, на основе чего сформирована гибридная модель прогнозирования, состоящая из двух автономных LSTM-ветвей нейроконтроллера. Предложенный подход сравнивался с единым LSTM-контроллером без кластеризации и с известными кластер-LSTM-методами при одинаковых вычислительных условиях. В качестве критерия использовалась ошибка прогнозирования MAPE. Результирующие значения отклонений расчетов двухпоточного нейроконтроллера от реальных данных на тестовом подмножестве по метрике MAPE приведены на рисунке 4.

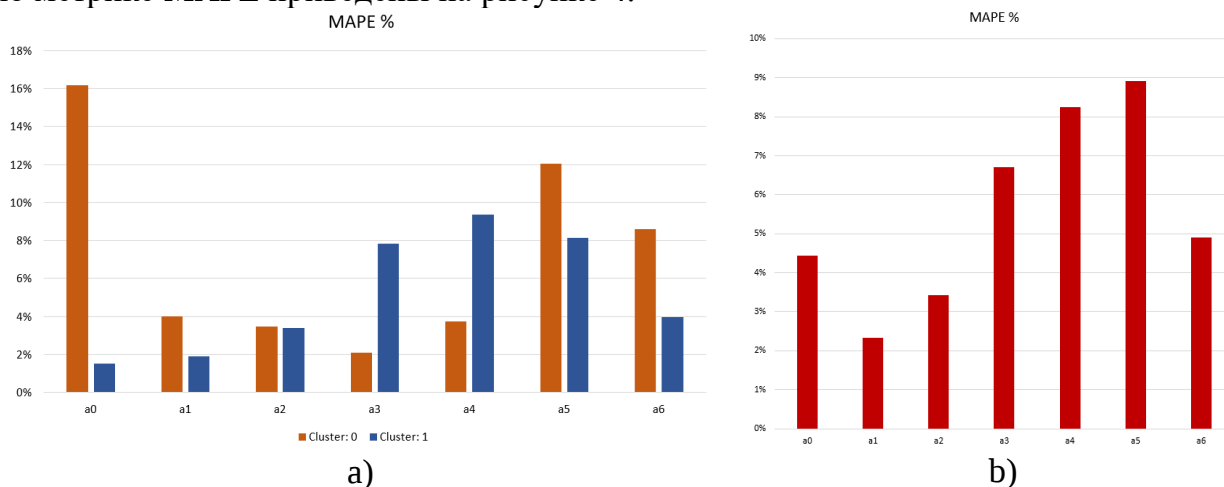


Рисунок 4 – Средние значения ошибок MAPE по управляющим коэффициентам $a_0 - a_6$: а) по каждому из выделенных кластеров; б) по всему датасету в целом

Показано, что разработанный нейроконтроллер обеспечивает среднюю ошибку 5,57% (точность 94,43%) при времени расчёта около 0,1 с на задачу, превосходя сравнимые методы по точности без увеличения вычислительных затрат.

Затем были проведены вычислительные эксперименты по прогнозированию уровня глюкозы в крови пациентов, больных сахарным диабетом 1 типа, на основе данных инвазивных сенсоров инсулиновых помп для смешанного временного ряда, состоящего из значений уровня глюкозы крови, активного инсулина в крови и активных углеводов. Временные ряды были сегментированы, описаны набором статистических и динамических признаков и кластеризованы согласно предложенному методу адаптивной кластеризации с выделением четырёх кластеров, соответствующих различным состояниям пациента. Затем была построена гибридная модель прогнозирования, состоящая из четырех отдельно обучаемых LSTM-веток. Результаты прогнозирования на тестовом смешанном временном ряде демонстрирует рисунок 5.

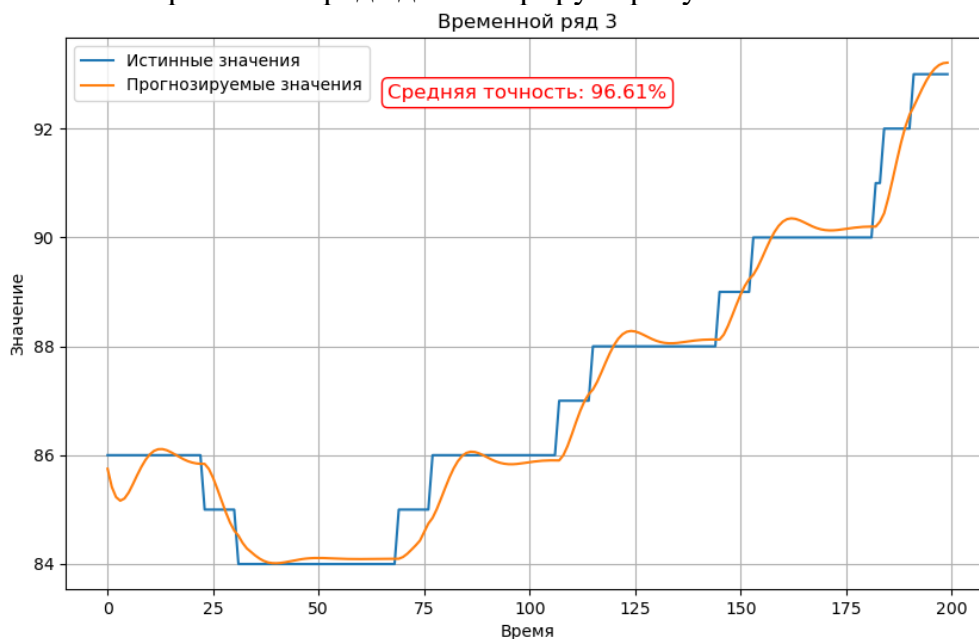


Рисунок 5 – Реальные и прогнозные значения уровня глюкозы для тестового временного ряда

Показано, что кластерный LSTM-подход обеспечивает среднюю точность прогнозирования до 95,23%, существенно превосходя LSTM без кластеризации и известные кластерные методы при сопоставимых вычислительных затратах (таблица 4).

Таблица 4 – Сравнение средней точности прогнозирования уровня глюкозы различными методами

Метод прогнозирования	Средняя точность, %
LSTM без кластеризации	81,72
LSTM с кластеризацией по Евклидовой метрике	83,46
LSTM с кластеризацией по методу ARIMA	91,69
Предложенный метод	95,23

Экспериментально подтверждена вычислительная эффективность предложенного программного комплекса, что делает возможным его применение в системах реального времени и в условиях ограниченных вычислительных ресурсов.

В заключении сформулированы основные научные и практические результаты работы, намечены направления перспективных исследований.

В приложении 1 представлены акты о внедрении и использовании результатов диссертационного исследования.

В приложении 2 представлены свидетельства о регистрации программ для ЭВМ.

ЗАКЛЮЧЕНИЕ

В диссертации на основе проведенных исследований решена актуальная научная задача, имеющая значение для развития и применения средств математического моделирования, численных методов и комплексов программ для повышения точности прогнозирования неоднородных временных рядов с минимальными вычислительными затратами в различных предметных областях. Обоснована необходимость разработки специальных методов и алгоритмов, обеспечивающих обработку временных рядов, характеризующихся высокой вариабельностью, шумами, пропущенными значениями и нестационарной динамикой. С учетом особенностей решаемой задачи сформулированы требования, разработаны методы и алгоритмы, а также программный комплекс для предварительной обработки, кластеризации и адаптивного прогнозирования временных рядов с указанными свойствами.

Разработанное в диссертации алгоритмическое обеспечение включает в себя следующие основные компоненты:

- метод и алгоритм адаптивной кусочно-линейной аппроксимации временных рядов с динамическим определением шага аппроксимации, обеспечивающий снижение размерности временных рядов при сохранении заданной точности;
- метод и алгоритм кластеризации временных рядов на основе вычисления временных характеристик, отражающих уровень, вариабельность, форму распределения и динамику данных, позволяющий формировать интерпретируемые и устойчивые кластеры;
- модель и алгоритм адаптивного прогнозирования временных рядов на основе LSTM-моделей с предварительной кластеризацией данных, обеспечивающие адаптивный выбор прогнозной модели.

Разработаны и реализованы программные модули, спроектированные в рамках единой программной архитектуры, реализующие предложенные методы и алгоритмы и обеспечивающие полный цикл обработки данных: от предобработки и кластеризации временных рядов до прогнозирования уровня глюкозы.

В качестве практической апробации проведены вычислительные эксперименты на реальных медицинских данных мониторинга уровня глюкозы у пациентов с сахарным диабетом, а также на данных о траекториях и характеристиках полета беспилотных летательных и гибридных летательно-подводных аппаратов. Результаты вычислительных экспериментов показали, что применение предварительной кластеризации и адаптивной обработки данных позволяет снизить ошибки прогнозирования и повысить устойчивость моделей к шумам с учетом индивидуальных и контекстно-зависимых особенностей данных. Экспериментально подтверждена эффективность предложенных методов адаптивной аппроксимации и кластеризации временных рядов на основе вычисления временных характеристик по сравнению с классическими подходами.

Проведено сравнение точности разработанных моделей прогнозирования с базовыми LSTM-моделями без кластеризации и с классическими методами анализа временных рядов. Показано, что использование разработанного подхода обеспечивает улучшение качества прогнозирования и снижение вычислительных затрат, что особенно

важно для применения в системах реального времени на маломощных портативных, в том числе бортовых, устройствах.

Таким образом, в диссертации решены все поставленные задачи, а цель работы, заключающаяся в разработке моделей, методов и программного обеспечения для повышения точности и адаптации прогнозов к индивидуальной динамике временных рядов, достигнута.

Рекомендуется применение разработанных методов и программных средств при создании интеллектуальных систем анализа и прогнозирования временных рядов различной природы.

Перспективы дальнейшей разработки темы связаны с развитием методов адаптивной кластеризации временных рядов, расширением архитектур нейросетевых моделей прогнозирования, а также применением разработанных методов для анализа больших потоков данных и интеграцией алгоритмов в распределенные вычислительные системы.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в журналах, входящих в перечень ведущих рецензируемых научных журналов и изданий, рекомендованных ВАК Министерства науки и высшего образования Российской Федерации:

1. Мингалиев, З. З. Нейросетевое прогнозирование гликемии у пациентов с сахарным диабетом на основе смешанных временных рядов с перспективой применения в составе интеллектуальной инсулиновой помпы / З. З. Мингалиев, С. В. Новикова // Современные информационные технологии и ИТ-образование. – 2021. – Т. 17, № 1. – С. 90-98.

2. Мингалиев, З. З. Метод расчета переменного шага для кусочно-линейной аппроксимации эмпирической нелинейной функции с заданной точностью / З. З. Мингалиев, С. В. Новикова, Г. В. Моисеев // Вестник Тверского государственного университета. Серия: Прикладная математика. – 2022. – № 3. – С. 35-48.

3. Мингалиев, З. З. Метод адаптивного прогнозирования данных, зависящих от времени, на основе LSTM-моделей с использованием кластеризации временных рядов / З. З. Мингалиев // Вестник Технологического университета. – 2025. – Т. 28, № 2. – С. 70-78.

4. Мингалиев, З. З. Программный комплекс для прогнозирования смешанных временных рядов с использованием кластеризации / З. З. Мингалиев // Вестник Технологического университета. – 2025. – Т. 28, № 3. – С. 103-110.

5. Мингалиев, З. З. Модель прогнозирования гликемии на основе смешанных временных рядов / З. З. Мингалиев, С. В. Новикова // Вестник Тверского государственного университета. Серия: Прикладная математика. – 2025. – № 1. – С. 67-87.

Статьи в журнале, включенном в базы цитирования Scopus

6. Mingaliyev, Z. Z. Dynamic step selection algorithm for piecewise linear approximation of complex control trajectories / L. Tan, L. Li, H. Su, S.V. Novikova, X. Zhang, Z. Z. Mingaliyev // Ocean Engineering. – 2023. – Vol. 280. – P. 114498.

Свидетельства о регистрации программ для ЭВМ

7. Мингалиев, З. З. Свидетельство о государственной регистрации программы для ЭВМ № 2022665809 Российская Федерация. Адаптивный модуль аппроксимации: № 2022664687: заявл. 08.08.2022: опубл. 22.08.2022 / З. З. Мингалиев; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ».

8. Мингалиев, З. З. Свидетельство о государственной регистрации программы для ЭВМ № 2025661819 Российская Федерация. Программа синтаксического анализа больших данных биологического мониторинга: заявл. 13.05.2025: опубл. 13.05.2025 / З. З. Мингалиев; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ».

9. Мингалиев, З. З. Свидетельство о государственной регистрации программы для ЭВМ № 2025661810 Российская Федерация. Программа кластеризации и адаптивного прогнозирования временных рядов на основе LSTM: заявл. 13.05.2025: опубл. 13.05.2025 / З. З. Мингалиев; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ».

10. Мингалиев, З. З. Свидетельство о государственной регистрации программы для ЭВМ № 2025661805 Российская Федерация. Программа анализа и прогнозирования смешанных временных рядов: заявл. 13.05.2025: опубл. 13.05.2025 / З. З. Мингалиев; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ».

Публикации в других научных изданиях и сборниках материалов конференций

11. Мингалиев, З. З. Программная реализация адаптивного модуля аппроксимации / З. З. Мингалиев // XXVI Туполевские чтения (школа молодых ученых): Материалы Международной молодёжной научной конференции. Сборник докладов, Казань, 09–10 ноября 2023 года. – Казань: ИП Сагиев А.Р., 2023. – С. 3138-3141.

12. Мингалиев, З. З. Интеллектуальное персонализированное нейросетевое прогнозирование гликемии у больных сахарным диабетом на основе смешанных временных рядов с перспективой внедрения в систему интеллектуальной инсулиновой помпы (ИИП) / С. В. Новикова, З. З. Мингалиев // Современные цифровые технологии: проблемы, решения, перспективы: национальная (с международным участием) научно-практическая конференция, Казань, 19–20 мая 2022 года. – Казань: Казанский государственный энергетический университет, 2022. – С. 3-11.

13. Мингалиев, З. З. Адаптивное прогнозирование временных рядов на основе кластеризации и нейросетевых моделей LSTM / З. З. Мингалиев, С. В. Новикова // Цифровые системы и модели: теория и практика проектирования, разработки и использования: Материалы международной научно-практической конференции, Казань, 10–11 апреля 2025 года. – Казань: Казанский государственный энергетический университет, 2025. – С. 1635-1639.

14. Мингалиев, З. З. Программный комплекс для адаптивного прогнозирования временных рядов / З. З. Мингалиев // Профессиональные коммуникации в научной среде – фактор обеспечения качества исследований: сборник материалов XV

Всероссийской научно-практической конференции, Альметьевск, 16 апреля 2025 года. – Санкт-Петербург: ООО Издательский дом «Сциентиа», 2025. – С. 623-626.

15. Мингалиев, З. З. Многофакторная нейросетевая модель краткосрочного прогнозирования гликемии у больных сахарным диабетом типа I / З. З. Мингалиев // Приоритетные направления инновационной деятельности в промышленности: сборник научных статей по итогам девятой международной научной конференции, Казань, 29–30 сентября 2020 года. Том Часть 2. – Казань: ООО «Конверт», 2020. – С. 30-34.

16. Мингалиев, З. З. Автоматический синтаксический анализатор больших данных биологического мониторинга / З. З. Мингалиев // Кулагинские чтения: техника и технологии производственных процессов: Материалы XX Международной научно-практической конференции. В 3-х частях, Чита, 30 ноября – 04 декабря 2020 года. Часть I. – Чита: Забайкальский государственный университет, 2020. – С. 33-37.