

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Уфимский университет науки и технологий»



На правах рукописи

ПАЛЬЧЕВСКИЙ Евгений Владимирович

МЕТОДЫ НЕЙРОСЕТЕВОЙ ОБРАБОТКИ БОЛЬШИХ
ТЕМПОРАЛЬНЫХ ДАННЫХ ДЛЯ ИНФОРМАЦИОННОЙ
ПОДДЕРЖКИ ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ
(НА ПРИМЕРЕ ЭЛЕКТРОЭНЕРГЕТИКИ)

Специальность 2.3.1. Системный анализ, управление и обработка информации,
статистика

Диссертация на соискание ученой степени кандидата технических наук

Научный руководитель:

доктор технических наук, доцент

Антонов Вячеслав Викторович

Уфа – 2024

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. АНАЛИЗ ОСОБЕННОСТЕЙ МЕТОДОВ И ПРОБЛЕМЫ ОБРАБОТКИ БОЛЬШИХ ТЕМПОРАЛЬНЫХ ДАННЫХ ПРИ ПРОГНОЗИРОВАНИИ ВРЕМЕННЫХ РЯДОВ И ФИЛЬТРАЦИИ СЕТЕВОГО ПОТОКА ДАННЫХ В ЭЛЕКТРОЭНЕРГЕТИКЕ.....	15
1.1. Актуальность и анализ проблемы управления и обработки данных в электроэнергетике	15
1.2. Классификация и анализ методов обработки больших темпоральных данных в электроэнергетике	24
1.3. Цель и задачи исследования.....	42
Выводы по главе 1	46
ГЛАВА 2. РАЗРАБОТКА МЕТОДОВ И МОДЕЛЕЙ ОБРАБОТКИ БОЛЬШИХ ТЕМПОРАЛЬНЫХ ДАННЫХ И ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ	48
2.1. Концептуальные положения решения задачи обработки и анализа больших темпоральных данных и прогнозирования временных рядов в электроэнергетике	49
2.2. Разработка метода и модели обработки сетевого потока данных с применением импульсной нейронной сети	53
2.3. Разработка метода и модели прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети.....	66
Выводы по главе 2	79
ГЛАВА 3. РАЗРАБОТКА ПРОТОТИПА СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ОБРАБОТКИ БОЛЬШИХ ТЕМПОРАЛЬНЫХ ДАННЫХ И ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ	81
3.1. Разработка системной модели и четырехкаскадной структурной схемы управления с включением прототипа СППР	81

3.2. Разработка информационной модели прототипа СППР.....	88
3.3. Требования по разработке прототипа СППР, архитектура, разработка и основной функционал	91
Выводы по главе 3	110
ГЛАВА 4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРЕДЛОЖЕННЫХ РЕШЕНИЙ НА ПРИМЕРЕ ПРЕДПРИЯТИЯ ЭНЕРГЕТИЧЕСКОГО СЕКТОРА.....	113
4.1. Оценка эффективности метода и модели обработки сетевого потока данных с применением импульсной нейронной сети на примере предприятия электроэнергетического сектора.....	115
4.2. Оценка эффективности метода и модели прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети на примере предприятий энергетического сектора.....	126
4.3. Оценка эффективности отдельных модулей прототипа СППР	145
4.4. Итоговое тестирование и оценка эффективности прототипа СППР	148
Выводы по главе 4	153
ЗАКЛЮЧЕНИЕ	156
СПИСОК ЛИТЕРАТУРЫ	161
ПРИЛОЖЕНИЕ А.....	183

ВВЕДЕНИЕ

В современном мире информатизация и цифровизация обуславливаются повсеместным внедрением специализированных информационных технологий во всех областях жизнедеятельности человека [1, 2]. Данное явление напрямую связано с таким глобальным научным и технологическим направлением как четвертая промышленная революция (4 IR), называемая в Европе «Индустрия 4.0» и плавно переходящая в «Индустрию 5.0», а в России – цифровая экономика [3, 4]. При этом одним из важнейших направлений в рамках цифровой экономики, является создание прототипа вычислительной системы в виде системы поддержки принятия решений (СППР) для выполнения широкого спектра задач, в том числе и прогнозирования.

Для достижения данной цели объединяются возможности киберфизических систем, использующих различные датчики, вычислительные сети (локальные и внешние глобальные) и модули, специализированные исполнительные механизмы с компьютерными моделями объекта (в основном математические), отражающие структуру и поведение (в том числе и обратную связь для самосинхронизации с состоянием реальных объектов) данного объекта [5–18].

Таким образом, появилась проблема изучения свойств сложной системы в том числе и с учетом анализа данных, полученных в результате работы информационной системой. При развитии информационных систем это очень актуально, т.к. не всегда возможно математически описать изучаемый процесс без анализа сетевого потока данных, и выявления некоторых характерных наблюдаемых величин в виде временного ряда. Это вызывает необходимость автоматизации объектов электрических сетей. При этом существует большое количество осложняющих факторов. Например, множество целей функционирования, взаимосвязь оцениваемых параметров, современные социально-экономические требования и т.д., которые приводят к увеличению объемов обрабатываемой информации, в том числе и носящей неопределенный и

зашумленный характер. Это обуславливает необходимость не только автоматизации информационных процессов, но и внедрения интеллектуальных систем поддержки принятия решения (СППР). Следовательно, одной из задач при реализации СППР для информационного сопровождения принятия решений в электроэнергетике является обработка в режиме реального времени поступающего потока данных из сети с одновременным прогнозом временных рядов и выявлением проблем взаимодействия с внешней средой. При этом СППР могут информационно взаимодействовать с другими системами под управлением системного программного обеспечения за счет разделения ресурсов как вычислительных мощностей, так и физических объектов.

В связи с этим **актуальность** диссертационного исследования обусловлена необходимостью разработки методов обработки больших темпоральных данных и прогнозирования временных рядов с целью повышения эффективности информационной поддержки принятия управленческих решений (например, проводить энергосберегающие мероприятия, начиная от контроля рабочих режимов и сроков проверки приборов учета электрической энергии и заканчивая реконструкцией электрических сетей, а также резервированием генерирующих мощностей электроэнергетических систем), что позволяет повысить точность экспертных заключений и уменьшает роль человеческого фактора при принятии решений в различных ситуациях.

СТЕПЕНЬ РАЗРАБОТАННОСТИ ТЕМЫ

Задачам обработки больших данных, в том числе и прогнозированию временных рядов в самых различных областях посвящено большое количество исследований как российских, так и зарубежных ученых, в которых широко обсуждаются проблемы повышения точности расчета прогнозных значений при большом прогнозном горизонте и вопросы обработки больших объемов информации, в том числе и с использованием нейросетевых технологий:

– задачами обработки больших данных занимались Бартоломей П.И. [19], Васильев В.И. [20], Вульфин А.М. [21], Гриценко Е.М. [22], Зегжда Д.П. [23], Тагирова К.Ф. [24], Четырбок П.В. [25], Devi B.S. [26], Peng T. [27] и др. В особенности было уделено внимание технологиям обработки пространственно-временных данных для их интеллектуального отслеживания, в том числе и в режиме реального времени, а также формальным абстрактным спецификациям не итеративных и итеративных подходов и обобщающих стратегий при обработке больших данных;

– задачами прогнозирования временных рядов занимались Грицай А.С. [28], Затонский А.В. [29], Надтока И.И. [30], Федоров И.В. [31], Филимонов Н.Б. [32], Juyong L. [33], Hadjout D. [34], Hai-Bao C. [35], Nishant J. [36], Xu R. [37], Yang M. [38], Yichen L. [39], Ying S. [40] и др. Углубленно рассматривались модели машинного и глубокого обучения при прогнозировании потребления электроэнергии, в том числе предлагалось и ансамблевое обучение с использованием трех разных моделей глубокого обучения: LSTM, GRU и TCN;

– задачами разработки систем поддержки принятия решений в различных областях занимались Аверченкова Е.Э. [41], Антонов В.В. [42], Вересников Г.С. [43], Гвоздев В.Е. [44], Гончаренко В.И. [45], Гузаиров М.Б. [46], Ильясов Б.Г. [47], Кулида И.П. [48], Маторин С.И. [49], Павлов С.В. [50], Сагитова Л.А. [51], Тиханычев О.В. [52], Юсупова Н.И. [53], Chinyere N.O. [54], Sutton R.T. [55] и др. Рассматривались задачи разработки и внедрения СППР в различных областях, в том числе и электроэнергетике с целью информационной поддержки принятия управленческих решений (например, проведение энергосберегающих мероприятий).

Перечень исследователей по вышеперечисленным направлениям может быть существенно расширен, но в большинстве вышеперечисленных работ авторов не приводится полное описание комплексного применения возможностей разных типов искусственных нейронных сетей (ИНС) для осуществления в режиме реального времени анализа и обработки поступающих данных, прогнозирования

временных рядов для предсказания возможных ситуаций, реакции систем на внешние воздействия с необходимой точностью выходных значений.

ОБЪЕКТОМ ИССЛЕДОВАНИЯ является процесс обработки больших данных и прогнозирования временных рядов для организации информационной поддержки принятия управленческих решений на примере электроэнергетики.

ПРЕДМЕТОМ ИССЛЕДОВАНИЯ являются модели и методы обработки информации в виде больших темпоральных данных и прогнозирования временных рядов на основе машинного и глубокого обучения (на примере электроэнергетики).

ЦЕЛЬ И ЗАДАЧИ ИССЛЕДОВАНИЯ

Целью диссертационной работы является повышение эффективности информационного сопровождения принятия управленческих решений в организации при проведении энергосберегающих мероприятий за счет нейросетевой обработки больших темпоральных данных с одновременным прогнозированием многомерных временных рядов¹.

Для достижения поставленной в диссертационной работе цели необходимо решить следующие задачи:

1. Разработать метод и модель обработки информации для формирования исходного датасета рекуррентной нейронной сети с целью прогнозирования значений потребления электроэнергии.

2. Разработать метод и модель прогнозирования значений потребления электроэнергии с применением модифицированной рекуррентной нейронной сети для информационной поддержки принятия управленческих решений.

3. Разработать структурную схему системы управления процессом поддержки принятия управленческих решений в электроэнергетике и архитектуру

¹ Многомерный временной ряд – система из нескольких одномерных рядов, где значения одного (целевого) зависят не только от предыдущих значений этого ряда, но и от значений дополнительного, или нескольких дополнительных

прототипа СППР на основе предложенных методов и моделей, реализованных в виде отдельных модулей.

4. Разработать прототип СППР в виде алгоритмического и программного обеспечения на основе нейросетевой обработки больших темпоральных данных и прогнозирования временных рядов для предприятия электроэнергетической отрасли.

5. Оценить эффективность предложенных решений (обработка больших объемов сетевого потока данных и прогнозирование потребления электроэнергии) на примере предприятия электроэнергетической отрасли.

Научная новизна и теоретическая значимость заключается в следующем:

1. Предложен метод обработки информации, включающий анализ и фильтрацию сетевого потока больших темпоральных данных на основе модели импульсной нейронной сети. Метод *отличается тем*, что на основе порядкового подхода следования импульсов (сетевой поток данных большого объема на транспортном и/или прикладном уровнях), модифицированного метода обучения обратного распространения ошибки и реализации возможности реагирования нейронов на внешние воздействия формируются специализированные правила обработки данных путем их нормализации, денормализации и кодирования, тем самым повышая скорость и качество обработки информации, *что позволяет* решать задачи информационной поддержки принятия решений при поступлении больших объемов данных через внешний сетевой интерфейс за счет формирования специализированных правил обработки информации.

2. Предложен метод нейросетевой обработки больших данных и прогнозирования временных рядов, основанный на модели рекуррентной нейронной сети и ее модифицированном методе обучения. Метод *отличается тем*, что на основе модифицированного метода обучения обратного распространения ошибки путем внедрения в него полученных экспериментально и формализованных коэффициентов скорости обучения ИНС и чувствительности, а

также модификации нейрона путем ввода специализированных коэффициентов учета аномалий (всплесков) во временных рядах повышается точность выходных прогнозных значений. *Это позволяет* повысить прогнозную точность (в зависимости от сезонности до 92–95%) временных рядов для решения задач информационной поддержки принятия управленческих решений.

3. Новизна структурной схемы, предложенной в рамках прототипа СППР, заключается в *четырёхкаскадном контуре управления* – четырех замкнутых контурах, в которых на основе предложенных методов и моделей объединены возможности импульсной нейронной сети (ИМНС) и рекуррентной нейронной сети (РНС) с одновременной обработкой и анализом больших данных, представленных в виде сетевого потока информации. *Это позволяет* управлять процессом информационной поддержки принятия управленческих решений на основе обработки больших разнородных и темпоральных данных и прогнозирования временных рядов.

ПРАКТИЧЕСКАЯ ЗНАЧИМОСТЬ

Заключается в том, что на основе разработанных методов создан прототип информационной системы интеллектуальной поддержки принятия решений для прогнозирования значений потребления электроэнергии в организациях энергетической отрасли, что позволяет повысить точность экспертных заключений и уменьшает роль человеческого фактора при принятии решений в различных ситуациях.

Практическая значимость результатов диссертационного исследования подтверждается их внедрением:

1. ООО «Башкирэнерго» – «Кумертауские электрические сети» (г. Кумертау). Внедрен метод нейросетевой обработки больших темпоральных данных на примере прогнозирования временных рядов с применением рекуррентной нейронной сети второго поколения на основе модифицированного метода обучения обратного распространения ошибки для информационной

поддержки принятия управленческих решений: необходимость в проведении энергосберегающих мероприятий, расчет необходимых энерго мощностей для обеспечения электроэнергией обслуживаемого района с учетом показателя эффективности. При этом одной из функций данного предприятия является оптимизация распределения электроэнергии с учетом бесперебойной работы потребителей и минимизация при этом избыточного резервирования энергетических мощностей. Более подробно данный процесс рассмотрен в разделе 3.1, а апробация прототипа СППР в рамках данного процесса – в разделе 4.4.1. Результаты апробации прототипа СППР показали, что разработанный метод прогнозирования потребления электроэнергии до 17,7% точнее относительно других ИНС и до 26,6% точнее относительно статистических/гибридных методов прогнозирования, а также до 95% быстрее относительно других методов обучения ИНС, что дает возможность расчета (коррекции) количества резервируемых мощностей и более оптимизированного проведения энергосберегающих мероприятий.

2. Павловская ГЭС Приуфимской ТЭЦ филиала ООО «БГК». При разработке научно-обоснованной методологии построения предиктивной аналитики и диагностики, основанной на комплексной оценке и анализе вибрационного состояния и режимных параметров гидроагрегата в виде структуры цифрового двойника гидроагрегата, объединенных в программный комплекс (систему) предиктивной аналитики и диагностики, внедрены: метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений; метод и модель обработки сетевого потока данных с применением импульсной нейронной сети. Это позволило: разработать формализованные правила объединения данных для мониторинга оборудования из разных источников (баз данных АСУТП и из файлов разных форматов); создать классификаторы дефектов и формализованных рекомендаций (правил) по их устранению или переключения от одного вида ремонта к другому; создать

прогнозные модели на основе применений нейронной сети и оценки остаточного ресурса на основе динамики развития ситуации потенциальных инцидентов (отказов контролируемых узлов агрегатов), как следствие, прогноз состояния оборудования; интегрировать прогнозные показатели с принятыми решениями для обеспечения интеллектуального информационного сопровождения.

3. ООО «Газпромнефть-Цифровые решения». Внедрены метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений. Данный метод нейросетевой обработки больших данных и прогнозирования временных рядов, основанный на модели рекуррентной нейронной сети и ее модифицированном методе обучения, внедрен при разработке программного обеспечения мониторинга оборудования на основе обработки больших потоков данных из разных источников и интеграции прогнозных показателей с принятыми решениями для обеспечения интеллектуального информационного сопровождения.

4. ФГБОУ ВО «Уфимский университет науки и технологий» – кафедра автоматизированных систем управления (г. Уфа). Основные результаты, в том числе и нейронные сети, диссертации применяются в учебном процессе для решения задачи моделирования, прогнозирования и управления, где требуется строить нечеткие системы из наборов цифровых данных по дисциплинам «Математическое моделирование сложных систем» и «Специальные информационно-аналитические системы». На сегодняшний день исследования на кафедре АСУ в рамках совершенствования (с научной и практической точек зрения) данной системы продолжаются.

ПОЛОЖЕНИЯ, ВЫНОСИМЫЕ НА ЗАЩИТУ

1. Метод и модель обработки информации в виде сетевого потока данных с применением импульсной нейронной сети для подготовки исходного датасета с

целью прогнозирования значений потребления электроэнергии (соответствует п. 4 паспорта специальности 2.3.1).

2. Метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений (соответствует п. 4 паспорта специальности 2.3.1).

3. Структурная схема системы управления процессом поддержки принятия управленческих решений в электроэнергетике на основе предложенных методов и моделей (соответствует п. 9 паспорта специальности 2.3.1).

4. Прототип СППР на основе нейросетевой обработки больших темпоральных данных и прогнозирования временных рядов (соответствует п. 10 паспорта специальности 2.3.1).

5. Оценка эффективности предложенных решений (обработка большого объема информации в виде сетевого потока данных и прогнозирование потребления электроэнергии) на примере предприятия электроэнергетической отрасли (соответствует п. 3 паспорта специальности 2.3.1).

ДОСТОВЕРНОСТЬ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

Достоверность научных результатов обусловлена корректным использованием методов системного анализа и математического моделирования, общей теории и проектирования систем, математического, компьютерного и нейросетевого моделирования в рамках обработки больших темпоральных данных на примере информации в виде сетевого потока данных и прогнозирования временных рядов в области электроэнергетики. Вводимые допущения по нагрузке на вычислительные ресурсы мотивировались проводимыми экспериментами и фактами, известными из практики.

АПРОБАЦИЯ РЕЗУЛЬТАТОВ

Основные материалы диссертационной работы докладывались и обсуждались на:

– международной научно-технической конференции «Перспективные информационные технологии», ПИТ-2017, Самара (2017);

– всероссийской научной конференции с международным участием «Информационные технологии интеллектуальной поддержки принятия решений», Уфа (2021).

Разработанный прототип СППР апробирован в рамках автоматического формирования исходного датасета и прогнозирования потребления электроэнергии в: ООО «Башкирэнерго» – «Кумертауские электрические сети» (г. Кумертау) при проведении энергосберегающих мероприятий и резервирования энерго мощностей; Павловской ГЭС Приуфимской ТЭЦ филиала ООО «БГК» при внедрении системы автоматизированной системы предиктивной аналитики и диагностики; ООО «Газпромнефть-Цифровые решения» для мониторинга энергетического оборудования; учебном процессе ФГБОУ ВО «УУНиТ» для решения задачи моделирования, прогнозирования и управления с применением технологий искусственного интеллекта.

СВЯЗЬ ТЕМЫ ИССЛЕДОВАНИЯ С НАУЧНЫМИ ПРОГРАММАМИ

Исследования в данном направлении выполнялись в период с 2015 по 2024 годы на базе Уфимского государственного авиационного технического университета и Уфимского университета науки и технологий (приказ от 08.07.2022 № 644 «О реорганизации Башкирского государственного университета и Уфимского государственного авиационного технического университета в форме слияния и создания нового вуза...»): в 2021-2022 гг. в рамках основной части государственных заданий ВУЗам № FEUE-2020-0007 на базе Уфимского государственного авиационного технического университета; в 2023-2024 годах на

базе Уфимского университета науки и технологий в рамках основной части государственных заданий ВУЗам № FEUE-2023-0007.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

По теме диссертационного исследования опубликованы 10 научных работ, из них 5 статей в изданиях из перечня, утвержденного ВАК (4 издания входят в перечень RSCI – K1, 1 издание – K2) России; 2 статьи, входящие в международную базу Scopus (Q3); 3 статьи входят в другие издания. Получено 6 свидетельств о государственной регистрации программ для ЭВМ.

ЛИЧНЫЙ ВКЛАД АВТОРА

Результаты исследования, представленные в диссертации, получены лично автором под руководством научного руководителя. Из публикаций, подготовленных в соавторстве, в диссертации представлен материал, принадлежащий лично автору.

СТРУКТУРА И ОБЪЕМ ДИССЕРТАЦИИ

Диссертация состоит из введения, четырех глав, заключения, списка использованных источников и приложений. Работа изложена на 194 страницах, в том числе: основной текст на 182 страницах, 33 таблицы, 39 рисунков, список использованных источников из 180 наименований на 22 страницах, 2 приложения на 12 страницах.

ГЛАВА 1. АНАЛИЗ ОСОБЕННОСТЕЙ МЕТОДОВ И ПРОБЛЕМЫ ОБРАБОТКИ БОЛЬШИХ ТЕМПОРАЛЬНЫХ ДАННЫХ ПРИ ПРОГНОЗИРОВАНИИ ВРЕМЕННЫХ РЯДОВ И ФИЛЬТРАЦИИ СЕТЕВОГО ПОТОКА ДАННЫХ В ЭЛЕКТРОЭНЕРГЕТИКЕ

В настоящей главе приводится анализ существующих методов и возникающих проблем при обработке больших темпоральных данных для поддержки принятия решений на примере обработки больших объемов сетевого потока данных и прогнозирования временных рядов в области электроэнергетики. Выделена значимость нейросетевых математических моделей при обработке и нормализации больших объемов данных, а также определены их особенности, преимущества и недостатки, что позволило определить место задачи поддержки принятия решения на основе интеллектуального анализа больших данных в задачах нейросетевой обработки информации. При этом увеличение количества наблюдаемых видов и типов данных затрудняет применение статистических, регрессионных и гибридных методов как для обработки данных в целом, так и для прогнозирования в частности. Таким образом, обозначенные проблемы определили цель и задачи диссертационного исследования.

1.1. Актуальность и анализ проблемы управления и обработки данных в электроэнергетике

В концепции «Цифровая трансформация 2030» [56], определяющей национальные цели и стратегические задачи развития Российской Федерации на период до 2030 года, указаны специализированные цели и задачи, являющиеся важным посылом для внедрения интеллектуальных информационных технологий управления и обработки больших данных во все сферы экономики Российской Федерации. Более того, согласно программе «Развитие энергетики» до 2035 года и доктрине «Энергетическая безопасность Российской Федерации» является необходимым развитие цифровых интеллектуальных технологий для внедрения их

в области энергетики, в том числе и в электроэнергетике. При этом основными вызовами для перехода к цифровой трансформации и цифровым интеллектуальным технологиям при обработке больших данных является повышение скорости генерации информации и увеличение объема данных. Из-за огромного объема и присущей этим данным темпоральности и разнородности, классические методы обработки данных, например, методы сравнительного анализа (горизонтальный, вертикальный и трендовый), не подходят для работы с ними [57]. В результате чего спрос на более быструю обработку больших данных ежедневно растет, и многие исследователи сосредоточились на модификации методов их обработки, сделав упор на точность при решении задач классификации, кластеризации и регрессии. Таким образом, возникает аспект, связанный с необходимостью развития программных и программно-аппаратных технологий анализа и обработки больших разнородных и темпоральных данных, в том числе и на основе интеграции информационно-вычислительных сетей и ИНС.

С начала 2000 года и по сегодняшний день количество исследований по моделированию и прогнозированию в электроэнергетике в Российской Федерации значительно возросло (таблица 1.1, рисунок 1.1), и их можно разделить на три основные группы [58]: *«Исследования причинно-следственной связи»*, *«Исследования взаимосвязей»*, *«Прогнозные исследования»*.

Таблица 1.1 – Количество публикаций в области прогнозирования потребления электроэнергии в 2020–2023 гг. в Российской Федерации

Год	Количество публикаций	Год	Количество публикаций	Год	Количество публикаций
2000	7	2008	132	2016	230
2001	8	2009	133	2017	276
2002	15	2010	132	2018	242
2003	14	2011	131	2019	264
2004	20	2012	231	2020	232
2005	48	2013	180	2021	220
2006	45	2014	230	2022	244
2007	86	2015	280	2023	203

Группы исследований: *«Исследования причинно-следственной связи»*, *«Исследования взаимосвязей»*, *«Прогнозные исследования»*

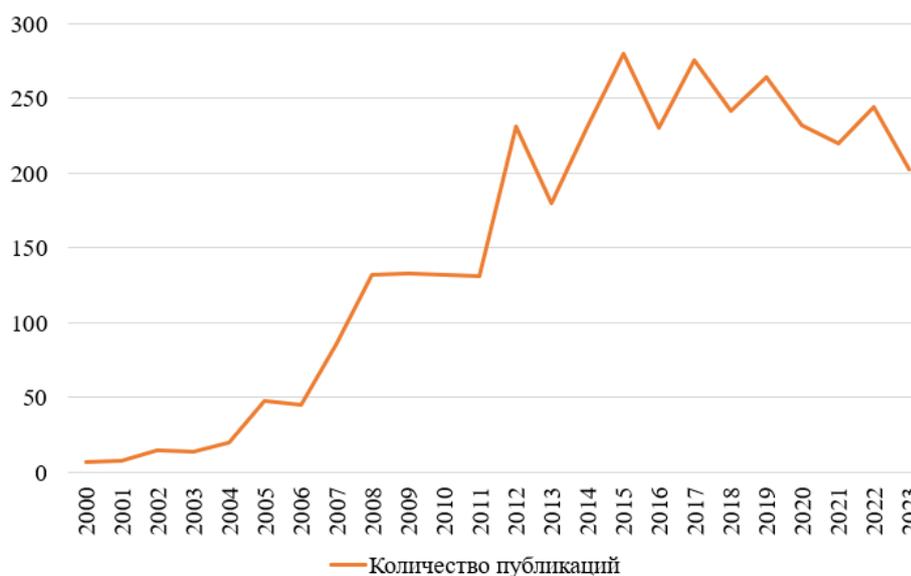


Рисунок 1.1 – Количество публикаций (по годам) в рамках прогнозирования потребления электроэнергии в Российской Федерации

В первой группе («*Исследования причинно-следственной связи*») основное внимание уделяется исследованию причинно-следственной связи между потреблением электроэнергии и экономическими составляющими, например, ВВП. Во второй группе («*Исследования взаимосвязей*») основное внимание уделяется выявлению взаимосвязи и величине ключевых взаимосвязей между экономическими переменными и потреблением энергии. А в третьей группе («*Прогнозные исследования*») цель состоит в том, чтобы спрогнозировать будущий спрос на электроэнергию, используя ряд различных подходов и моделей (таблица 1.2).

Таблица 1.2 – Краткое изложение обзора литературы по исследованиям прогнозирования потребления электроэнергии и проблемы управления в электроэнергетике Российской Федерации

Год	Ссылка	Фокус исследования	Модель/метод прогнозирования, а также их вид	Краткое содержание
2000	[59]	Устойчивое развитие электроэнергетики и проблема управления в энергетических предприятиях	Регрессионные, статистические модели	Рассмотрено прогнозирование потребления электроэнергии в рамках устойчивого развития электроэнергетики в Российской Федерации
2001	[60]	Взаимоотношениям топливно-энергетического сектора с экономикой Российской Федерации	Полиномиальная регрессия, статистические модели	Сделан упор на прогнозирование потребления топливно-энергетических ресурсов, в том числе и электроэнергии
2002	[61]	Прогнозирование удельного расхода электроэнергии при производстве водорода	Проприетарная математическая модель	Разработка новой математической модели для прогнозирования потребления электроэнергии
2003	[62]	Снижение потерь электроэнергии в электрических сетях	Статистические модели и методы в рамках проприетарной системы АСКУЭ-БЫТ	Прогнозирование и учет потребления электроэнергии на основе системы АСКУЭ-БЫТ
2004	[63]	Энерготехнологический метод прогнозирования рационального энергопотребления и рассмотрение проблемы управления в энергетических предприятиях	Проприетарная энерготехнологическая модель	Предложен энерготехнологический метод прогнозирования рационального энергопотребления
2005	[64]	Анализ поведения энергетической отрасли в разные периоды развития экономики	Динамическая модель в виде линейной регрессии, статистические методы	Прогнозирование производства, потребления, экспорта, импорта и потерь электроэнергии

Таблица 1.2 – Продолжение

2007	[2]	Прогнозирование потребления электроэнергии	Рекуррентная нейронная сеть, нейросетевая модель	С использованием системного подхода разработан концептуальный базис построения НСМ, имеющей высокие прогностические свойства в условиях сильной зашумленности базы данных
2008	[65]	Прогнозирование потребления электроэнергии цеха агломерации	Множественная нелинейная регрессия, статистические модели	Разработан статистический метод прогнозирования потребления электроэнергии цехом агломерации горно-обогатительного комбината
2009	[66]	Прогнозирование потребления энергоресурсов в учреждениях бюджетной сферы	Полиномиальная регрессия, статистические модели	Применение модели полиномиальной регрессии при краткосрочном прогнозировании потребления электроэнергии
2010	[67]	Прогнозируются экономические показатели перспектив развития угледобычи и электроэнергетики, а также рассматривается проблема управления в энергетических предприятиях	Полиномиальная регрессия, статистические модели	Приведены фактические и прогнозные цены на уголь на мировом и российском рынках, а также в Восточно-Сибирском регионе в сопоставлении с объемами его добычи. Дана оценка возможности использования прогнозных экономических показателей для планирования угледобычи
2010	[68]	Прогнозирование потребления электроэнергии	Статистические модели и методы: кусочно-линейная функция, полином 4 порядка, нейросетевая модель: многослойный персептрон	Прогнозирование потребления электроэнергии различными методами на промышленных предприятиях

Таблица 1.2 – Продолжение

2011	[69]	Нормирование и прогнозирование потребления электроэнергии	Экспоненциальное сглаживание, статистические модели	Разработка системы нормирования и прогнозирования потребления электроэнергии для Магнитогорского металлургического комбината
2012	[70]	Краткосрочное оперативное прогнозирование потребления электроэнергии	Многослойный перцептрон, нейросетевая модель	Разработка многослойного перцептрона для краткосрочного и оперативного прогнозирования потребления электроэнергии на примере Самарской энергосистемы
2013	[71]	Попытка прогнозирования потребления электроэнергии с большим прогнозным горизонтом	Многослойный перцептрон, нейросетевая модель	Представлен подход, использующий теорию искусственных нейронных сетей для прогнозирования потребления электрической энергии, при условии минимизации ошибки прогнозирования на длительном интервале упреждения
2014	[72]	Описание методики построения системы прогнозирования электрической нагрузки с использованием методов вейвлет-анализа и гибридных нейронечетких систем	Интеграция вейвлет-анализа (статистические методы) и ANFIS (нейронная сеть), гибридные методы	Описание этапов разработки системы прогнозирования потребления электроэнергии

Таблица 1.2 – Продолжение

2015	[73]	Поиск оптимальной нейронной сети для краткосрочного прогнозирования потребления электроэнергии	Нейросетевые модели прогнозирования: линейная; многослойный персептрон; двухслойная Элмана; рекуррентная	Выбор оптимальной ИНС с точки зрения ее архитектуры и параметров для краткосрочного прогнозирования потребления электроэнергии
2016	[74]	Применение моделей машинного обучения для прогнозирования потребления электроэнергии	Дерево решений, статистические методы	Прогнозирование потребления электроэнергии на примере агента ветроустановки
2017	[75]	Проблема суточного прогнозирования потребления электроэнергии в Московском регионе (Москва и Московская область)	Многослойный персептрон, нейросетевая модель	Решение проблемы краткосрочного (суточного) прогнозирования потребления электроэнергии для последующего прогнозирования суточных объемов пассажирских перевозок в Московском метрополитене
2018	[76]	Прогнозирование потребления электроэнергии в рамках Единой энергетической системы	Полиномиальная регрессия, статистические методы	Формируется прогноз с прогнозирования потребления электроэнергии для «Единой энергетической системы» с учетом метеопказателей
2019	[77]	Прогнозирование потребления электроэнергии при помощи многослойного персептрона в Matlab	Рекуррентная нейронная сеть, нейросетевые модели	Долгосрочное нейросетевое прогнозирование потребления электроэнергии Барановичской дистанции электроснабжения Белорусской железной дороги

Таблица 1.2 – Окончание

2020	[78]	Многоуровневое прогнозирование и нормирование потребления электроэнергии	Полиномиальная регрессия, статистические методы	Многоуровневое прогнозирование потребления электроэнергии в рамках металлургического предприятия
2021	[79]	Прогнозирование потребления электроэнергии промышленного предприятия	ARIMA, статистические методы	Краткосрочное прогнозирование потребления электроэнергии на промышленном предприятии
2022	[80]	Прогнозирование потребления электроэнергии промышленного предприятия методами машинного обучения	Статистические модели: градиентный бустинг, случайный лес	Применение методов машинного обучения при прогнозировании потребления электроэнергии в рамках промышленного предприятия
2023	[81]	Нейросетевое прогнозирование потребления электроэнергии	Многослойный персептрон, нейросетевые модели	Применение многослойного персептрона при краткосрочном прогнозировании потребления электроэнергии

Таким образом, исходя из таблиц 1.1–1.2 видно, что в большинстве случаев в Российской Федерации для прогнозирования потребления электроэнергии преобладают статистические модели (пункт 1.2.1), при этом с разной периодичностью применяются как гибридные (пункт 1.2.3) и нейросетевые (пункт 1.2.4). Более того, нейросетевые методы прогнозирования потребления электроэнергии начали набирать популярность только ближе к 2020 году, что говорит о необходимости их дальнейшего развития. Однако существенное количество работ направлено на краткосрочное прогнозирование потребления электроэнергии, что напрямую влияет на долгосрочное резервирование энерго мощностей, как следствие, вынуждает энергопредприятия выделять

дополнительные финансовые средства. При этом проблемой долгосрочного прогнозирования является не только точность прогнозных значений, но и обработка больших данных, в том числе и их предобработка. Более того, в электроэнергетике с точки зрения обеспечения качественной обработки больших разнородных и темпоральных данных² одной из основных проблем является их классификация. Классификация представляет собой контролируемое машинное обучение, решающее проблему построения модели из правильно классифицированного набора данных. Таким образом, исходя из анализа литературных источников, в частности из таблиц 1.1-1.2, и несмотря на то, что в литературе на сегодняшний день есть огромное множество алгоритмов (методов) и моделей классификации и регрессии, в том числе и прогнозирования [82–84]: статистические (регрессионные) [85–86] (пункт 1.2.1), модели Грея [87–92] (пункт 1.2.2), гибридные модели [93–102] (пункт 1.2.3) и нейросетевые [103–121] (пункт 1.2.4), их комплексное использование в самых различных областях не исключает возникновения сложностей и аномалий при обработке больших разнородных и темпоральных данных для различных выборок. Таким образом, возникают две нерешенные технические задачи:

1. Автоматизированная предобработка больших темпоральных данных, получаемых со smart-датчиков и счетчиков, с целью повышения точности прогнозных значений.

2. Долгосрочное прогнозирование (до 30 суток вперед) потребления электроэнергии с высокой точностью (92–97%).

Одним из способов решения этих задач является:

1. Разработка нейросетевых моделей долгосрочного прогнозирования потребления электроэнергии с автоматизированной предобработкой больших разнородных и темпоральных данных.

² Дата измерения, № счетчика, потребляемая электроэнергия, температура воздуха, количество осадков, выходной день или нет, атмосферное давление и сезон

2. На основе разработанных математических моделей нейронных сетей – создание единой системы обработки больших разнородных данных с их последующим преобразованием в темпоральные на примере обработки сетевого потока данных со smart-датчиков и счетчиков электрических подстанций и прогнозирования временных рядов потребления электроэнергии для информационного сопровождения поддержки принятия управленческих решений.

1.2. Классификация и анализ методов обработки больших темпоральных данных в электроэнергетике

К настоящему моменту времени разработано большое количество моделей и методов обработки больших разнородных и темпоральных данных и прогнозирования временных рядов, а также четко определилась тенденция их комбинирования с целью получения лучших характеристик, в том числе и прогнозных значений.

В данном разделе представлены классификация и анализ методов и моделей обработки больших разнородных и темпоральных данных в электроэнергетике как инструментов обработки больших данных и прогнозирования временных рядов. При этом рассмотрены некоторые аспекты интеллектуального анализа различных данных, характеризующихся различной интенсивностью волатильности значений, форматов, неоднозначности, недостаточной полноты значений. Обозначены проблемы выявления аномалий во временных рядах, в том числе и при обработке неоднородных данных.

1.2.1. Статистические методы обработки больших данных с возможностью прогнозирования временных рядов в электроэнергетике

Существует большое количество статистических методов обработки больших данных и прогнозирования временных рядов в электроэнергетике, в том

числе ARIMA, ARMA, SVR, полиномиальная регрессия, авторегрессия, VECM, модели экспоненциального сглаживания. При этом среди перечисленных методов прогнозирования временных рядов с последующей обработкой больших данных наибольшее распространение получил метод ARIMA [122–124], включающий в себя модели авторегрессии и скользящего среднего, а также их различные модификации. Основными причинами (преимуществами) использования данных статистических методов являются простота в реализации, возможность обработки больших структурированных данных и доступность среди готовых решений в виде специализированных библиотек, написанных на различных языках программирования. Например, в языке высокоуровневого программирования Python примером реализации ARIMA является библиотека «Pyramid-arima» [125].

Стандартная математическая модель ARIMA может быть представлена следующей формулой:

$$EE_t = \alpha + \beta_1 EE_{t-1} + \beta_2 EE_{t-2} + \dots + \beta_p EE_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_q \varepsilon_{t-q}, \quad (1.1)$$

где: EE_t – прогнозируемое значение потребления электроэнергии; t – коэффициент задержки временного ряда; $EE_{t-1}, EE_{t-2}, \dots, EE_{t-p}$ – запаздывание (сдвиг) каждого временного ряда, подаваемого на вход для прогнозирования; ε_t – коэффициент ошибки прогнозирования; $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ – коэффициент ошибки прогнозирования с учетом задержки временного ряда; $\beta_1, \beta_2, \dots, \beta_p$ – лаговый оператор сдвига, позволяющий получить значения временного ряда на основании предыдущих элементов; ϕ_1, \dots, ϕ_q – авторегрессионный многочлен; p – порядок авторегрессии, позволяющий добавить предыдущие значения временного ряда; q – порядок скользящего среднего, позволяющий установить погрешность модели как линейную комбинацию наблюдавшихся ранее значений ошибок $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$.

В исследовании [126] приведены результаты прогнозирования краткосрочной нагрузки на электроэнергетическое оборудование с использованием гибридной регрессии опорных векторов и моделированием случайного леса. В рамках выполненной работы [127] ARIMA и SVR рассмотрены

в качестве инструментов для одновременного прогнозирования нескольких временных рядов, что делает разработанный метод более гибким в отношении ежедневного, еженедельного и годового прогнозирования. В исследовании [128] разработана новая гибридная модель прогнозирования EMD-SVR-PSO-AR-GARCH для прогнозирования временных рядов. Преимущество данной модели заключается в сочетании в себе нескольких подходов к машинному обучению: EMD, SVR и PSO. Это позволяет повысить точность краткосрочного прогнозирования. В [129] было проведено исследование, заключающееся в сравнении исходных и модифицированных моделей SVR, и LSTM. Результаты показали, что относительно других моделей у SVR самая большая среднеквадратичная ошибка (RMSE). В [130] выполнено исследование, посвященное статистическому прогнозированию временных рядов, в том числе и методом ARIMA. В [131] для прогнозирования потребления электроэнергии используется модель глубокого обучения ARIMAX-GARCH.

Модель ARIMA и ее модификации очень хорошо справляются с краткосрочным прогнозированием (на несколько суток) значений временных рядов, но не являются достаточно точными при раннем (заблаговременном) прогнозировании по причине достаточно сложного расчета прогнозных значений из-за наличия большого количества аномалий во входных параметрах объекта. Ограничение работы метода (моделей) ARIMA представлено на рисунке 1.2.

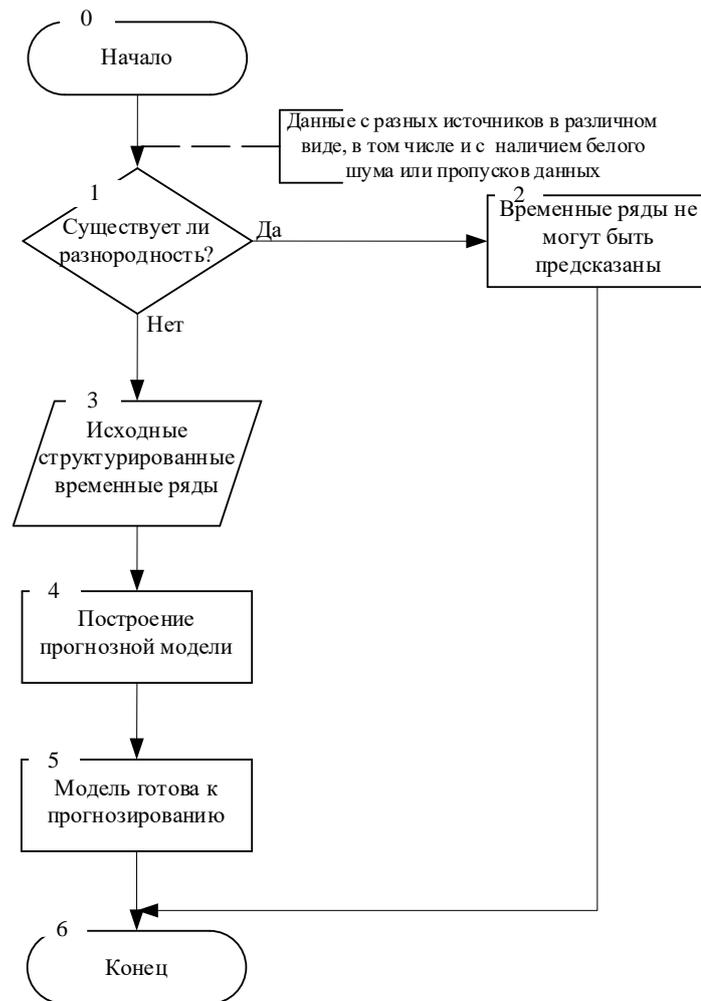


Рисунок 1.2 – Схема алгоритма, показывающая работу и ограничения моделей ARIMA

Таким образом, исходя из стандартной работы ARIMA (рисунок 1.2) можно отметить, что данная модель не имеет возможности обработки больших разнородных данных (например, наличие белого шума – пропусков данных), как следствие, необходимо выполнить разностную аппроксимацию входных параметров объекта (временного ряда).

Основными недостатками вышеприведенных методов являются:

1. Невозможность корректного расчета прогнозных значений временных рядов при высоком значении горизонта прогнозирования, т.е. при долгосрочном прогнозировании, как правило, больше 4 дней.

2. В случае предоставления на вход разнородных данных прогнозирование становится просто невозможным: нет четкой структурированности и

фаззификации входного массива данных, а для статистических методов это является основополагающим фактором обработки информации.

Отдельно хотелось бы выделить *модели экспоненциального сглаживания из-за их большого количества и разнообразия*. К основным моделям экспоненциального сглаживания относятся: классическая (модель Брауна, линейный тренд); аддитивная сезонность с демпфированным, линейным, исключенным и экспоненциальным трендами; модель Хольта (линейный тренд) и Хольта-Уинтерса.

Классическая модель экспоненциального сглаживания. Данная модель предназначена для временных рядов, в которых отсутствуют тренд (тенденция изменения) и сезонность. Основная суть заключается в экспоненциальных средних, благодаря чему у прогнозной модели появляется свойство последовательной адаптации к аномалиям во временных рядах, при этом без значительного реагирования на случайные отклонения. Данная модель описывается следующей системой уравнений:

$$\begin{cases} E_1 = EE_0 \\ E_t = \alpha \cdot EE_{t-1} + (1 - \alpha) \cdot E_{t-1} \end{cases}, \quad (1.2)$$

где: E_1 – исходный временной ряд; EE_0 – начальное фактическое наблюдение потребляемой электроэнергии в исходном временном ряде; EE_{t-1} – фактическое наблюдение потребляемой электроэнергии в момент времени $t-1$; E_t – значение экспоненциального среднего в момент времени t ; α – коэффициент сглаживания, изменяющийся в диапазоне от 0 до 1. Таким образом, при последовательном использовании (1.2), E_t можно выразить через значения временного ряда потребляемой электроэнергии:

$$E_t = \alpha \cdot \sum_{i=0}^{\infty} (1 - \alpha)^i \cdot EE_{t-i}, \quad (1.3)$$

где i – временной индекс удаленности наблюдения от настоящего момента t . При этом экспоненциальное уменьшение весов зависит от глубины наблюдений

относительно t . Для использования краткосрочных прогнозов используется следующая формула:

$$E_t = \alpha \cdot d_1 + (1 - \alpha) \cdot d_2, \quad (1.4)$$

где: d_1 и d_2 являются коэффициентами умножения, изменяющиеся в диапазоне от 0 до 1 и определяющиеся в зависимости от выбранной модели сглаживания, для простого экспоненциального сглаживания могут быть определены из формулы $d_1 = EE_t$ и $d_2 = E_{t-1}$; $1 - \alpha$ представляет собой эквивалентную постоянную сглаживания.

Исходя из эмпирического анализа [132] модели простого экспоненциального сглаживания было выявлено, что с ее помощью нецелесообразно:

1. Осуществлять долгосрочные прогнозы (больше 4-х суток) потребления электроэнергии из-за существенных отклонений прогнозных от фактических значений.

2. Выполнять любой прогноз для данных, имеющих ярко выраженную тенденцию (тренд).

В качестве преимущества необходимо отметить возможность наиболее простой, в том числе и программной реализации метода прогнозирования временных рядов.

Аддитивная сезонность с демпфированным трендом. Суть данной модели, в отличие от классической, заключается в наличии демпфированного (затухающего) тренда и аддитивной сезонности, выражаемой в виде аддитивной компоненты, что позволяет повысить прогнозное значение потребления электроэнергии. Модель описывается следующей системой уравнений:

$$\begin{cases} E_t = \frac{\left(\frac{1}{\alpha}\right) \cdot (EE_k^t - EE_1^t)}{|(k-1) \cdot p|}, \\ S_t = EE_k^t - p * E_t / 2 \end{cases}, \quad (1.5)$$

где:

α – коэффициент сглаживания, изменяющийся в диапазоне от 0 до 1;

EE_1^t – среднее значение потребляемой электроэнергии на первом сезонном цикле;

k – число полных сезонных циклов;

EE_k^t – среднее значение потребляемой электроэнергии на последнем сезонном цикле;

S_t – сезонная вариация среднего значения потребления электроэнергии;

E_t – значение тренда потребления электроэнергии

t – порядковый номер анализируемого временного интервала

p – длина сезонного цикла, целое число, большее или равное 2. Например, для анализируемого месячного цикла принимает значение равное 12.

Данная модель подходит для краткосрочного прогнозирования временных рядов, в которых наблюдаются линейный тренд и не изменяющаяся с течением времени сезонная вариация.

Детрендированная аддитивная сезонность. Отчасти данная модель является похожей на классическую модель экспоненциального сглаживания. При этом дополнительно в каждом прогнозе учитывается аддитивный сезонный тренд, выражаемый в виде аддитивной сезонной компоненты. Данная модель (1.6) хорошо подходит под прогнозирование временных рядов, в которых наблюдается минимальное количество аномалий [178].

$$\begin{cases} E_1 = EE_0 \\ E_t = \frac{(EE_k - EE_1)}{p} \end{cases} \quad (1.6)$$

Модель экспоненциального сглаживания с детрендированной с аддитивной сезонностью подходит для краткосрочного прогнозирования структурированных временных рядов, в которых отсутствуют тренд и сезонность.

Данная модель подходит для краткосрочного прогнозирования временных рядов, в которых наблюдается затухание тренда и отсутствует сезонность [179].

Аддитивная сезонность с линейным трендом. В данной модели учитываются как линейный тренд, так и аддитивная сезонная компонента. В случае, если нет резких колебаний во временном ряде, краткосрочный прогноз

получается относительно точным и описывается следующей системой уравнений (при обозначениях переменных по формуле 1.5):

$$\begin{cases} E_t = \frac{(EE_k^t - EE_1^t)}{|(k-1) \cdot p|} \\ S_t = EE_k^t - E_t / 2 \end{cases} \quad (1.7)$$

Данная модель подходит для прогнозирования временных рядов с постоянно изменяющимся трендом и сезонностью.

Модель Хольта представляет динамику временного ряда как линейную зависимость с изменяющимися параметрами при каждом прогнозе [178]. Данная модель используется в том случае, если во временном ряде имеется линейный тренд и отсутствует сезонность. При этом начальные прогнозные значения рассчитываются по методу наименьших квадратов. Данная модель описывается следующей системой уравнений:

$$\begin{cases} E_p(t, k) = E_0(t) + k \cdot E_1(t) \\ e(t+1) = E_p(t+1, 1) - E_p(t, 1) \\ E_0(t) = E_0(t-1) + E_1(t-1) + \alpha \cdot \exp(t) \end{cases} \quad (1.8)$$

где: $E_p(t, k)$ – прогнозная оценка p -го временного ряда потребления электроэнергии на горизонт k ; $E_0(t)$ – оценка t -го значения временного ряда потребления электроэнергии; $E_1(t)$ – оценка текущего изменения t -го значения потребления электроэнергии во временном ряде; $e(t+1)$ – отклонения прогнозируемого значения от реального при $k=1$; α – коэффициент адаптации (сглаживания), изменяющийся в диапазоне от 0 до 1.

У данной модели есть преимущество в гибкости выбора соотношений отслеживания уровня и наклона тренда во временном ряде, что повышает прогнозную точность при краткосрочном прогнозировании. Основным недостатком является невысокая точность при долгосрочном прогнозировании.

Модель Хольта-Уинтерса. Основным отличием от классической модели Хольта является расширение до трех параметров в рамках экспоненциального

сглаживания, что и является основным преимуществом. С применением данных параметров и будет строиться прогноз. Основным недостатком является невысокая точность при долгосрочном прогнозировании [132].

Модель описывается следующим образом: расчет экспоненциально-сглаженного ряда (1.9); определение значения тренда (1.10); оценка сезонности (1.11); расчет прогнозных значений потребления электроэнергии (1.12).

$$EE_t = \frac{\alpha \cdot E_{ts}}{E_{t-s} + (1 - \alpha) \cdot (EE_{t-1} + E_{t-1})}. \quad (1.9)$$

$$E_t = b \cdot (EE_t - EE_{t-1}) + (1 - b) \cdot E_{t-1}. \quad (1.10)$$

$$E_t = \frac{q \cdot E_{ts}}{EE_t + (1 - q) \cdot E_{t-s}}. \quad (1.11)$$

$$\hat{E}_{t+p} = (EE_{t+p} \cdot E_t) \cdot E_{t-s+p}. \quad (1.12)$$

В формулах (1.9–1.12): EE_t – сглаженная величина ряда на текущий период; α – коэффициент адаптации (сглаживания), изменяющийся в диапазоне от 0 до 1; b – коэффициент сглаживания тренда, изменяющийся в диапазоне от 0 до 1; q – коэффициент сглаживания сезонности, изменяющийся в диапазоне от 0 до 1; p – порядковый номер периода, на который производится прогнозирование; E_{t-s} – коэффициент сезонности предыдущего периода; E_{ts} – текущее значение временного ряда потребления электроэнергии; EE_{t-1} – сглаженная величина ряда за предыдущий период; E_{t-1} – значение тренда потребления электроэнергии за предыдущий период; \hat{E}_{t+p} – прогнозное значение на p периодов.

1.2.2. Модели Грея для обработки больших данных с возможностью прогнозирования временных рядов

Данные модели (схема алгоритма работы приведена на рисунке 1.3) хорошо подходят для прогнозирования нелинейных временных рядов (т.е. временных рядов с нелинейным трендом) потребления электроэнергии и относятся к

нестатистическим методам прогнозирования. Наибольшая эффективность данных моделей проявляется в высокой точности краткосрочного прогнозирования в условиях наличия белого шума – недостаточного количества ретроспективных данных, что является основным преимуществом данной модели [133]. Данная модель и ее модификации являются достаточно распространенными при краткосрочном и раннем (заблаговременном) прогнозировании временных рядов с учетом обработки больших данных [134].

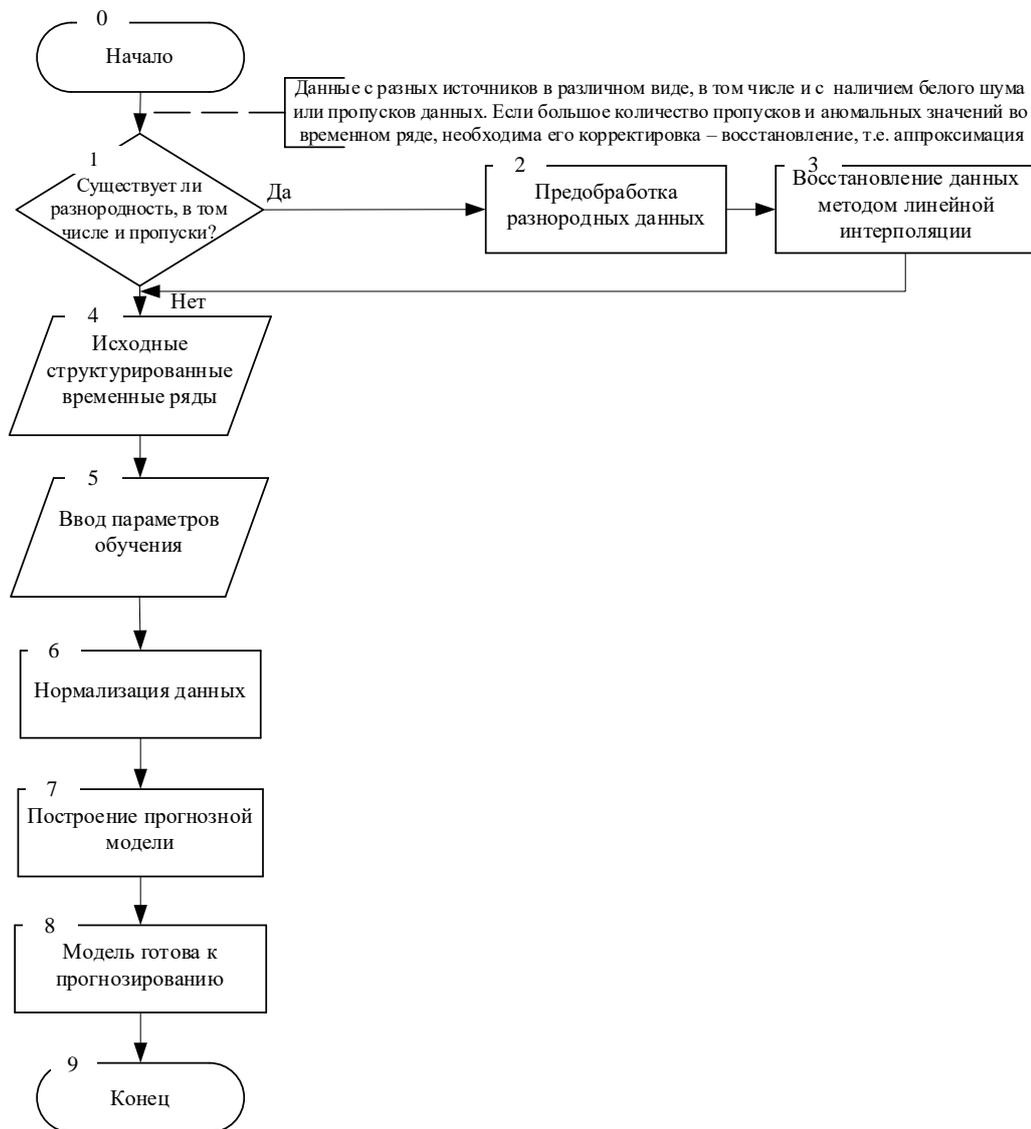


Рисунок 1.3 – Схема алгоритма стандартной предобработки данных в рамках модели Грея

Стандартная математическая модель Грея представлена в формулах (1.13–1.14):

$$TS = (EE(1), EE(2), \dots, EE(n)), \quad (1.13)$$

где TS – подаваемый на вход модели временной ряд, $EE(i)$ (при этом $i=1, \dots, n$) – значение потребления электроэнергии во временном ряде, n – порядковый номер значения потребления электроэнергии во временном ряде. Исходя из этого, добавив порядковый номер временного ряда, можно определить модель Грея дифференциальным уравнением первого порядка:

$$EE_{forecast}(k) = \frac{dTS(k)}{dk} + \alpha TS(k), \quad (1.14)$$

где $EE_{forecast}(k)$ – прогнозируемый временной ряд значений потребления электроэнергии, k – порядковый номер временного ряда, подаваемого для прогнозирования. Таким образом, исходя из (1.14) по методу наименьших квадратов, и применяя дискретное решение $EE(k)$ дифференциального уравнения (1.15) можем рассчитать предсказываемый временной ряд:

$$EE_{forecast}(1) = EE(1), \quad EE_{forecast}(k) = EE(k) - EE(k-1), \text{ где } k = 2, 3, \dots, n. \quad (1.15)$$

При этом данные модели зачастую применяются на практике при краткосрочном прогнозировании. Например, в работах [133, 134] для прогнозирования интенсивности потребления электроэнергии используются классическая модель Грея и интегрированная с методом наименьших квадратов модель Грея в сочетании с технологией вычисления опорных векторов для прогнозирования временных рядов в области электроэнергетики. В исследовании [135] авторы предложили новую несмещенную на основе модели Грея модель Бернулли со специализированным алгоритмом оптимизации и методом обработки больших данных для прогнозирования временных рядов. В [136] авторы предложили новую сезонную модель на основе модели Грея с использованием алгоритмов обработки больших данных для прогнозирования временных рядов. В [137] разработали новую модель Грея для прогнозирования сезонных временных

рядов со структурными разрывами и обработки больших данных на примере потребления количества электроэнергии. На основе классической модели Грея реализовали новую конформную дробно-неоднородную дискретную модель Грея CFNDGM с целью прогнозирования временных рядов и обработки больших данных [138].

Несмотря на распространение моделей Грея и их успешное применение в самых различных областях, у них зачастую наблюдаются неточные результаты при краткосрочном (за исключением случаев, когда данные уже структурированы и найдена функциональная зависимость) и раннем (заблаговременном) прогнозированиях. Основные причины: недостаточно проработанная функция временного отклика TRF, основанная на гипотезе о фиксированном промежутке времени, что ограничивает адаптивность данных моделей к различным ситуациям; невозможность обработки больших разнородных данных в силу тривиальности своих исходных алгоритмов.

Таким образом, исходя из ограничений стандартной модели Грея можно отметить, что данная модель может восстанавливать данные, но не имеет возможности обработки больших разнородных данных (например, в виде сетевого потока данных), как следствие, необходимо выполнить разностную аппроксимацию входных параметров объекта (временного ряда), как и у статистических методов обработки больших данных и прогнозирования временных рядов.

1.2.3. Гибридные методы обработки больших данных с возможностью прогнозирования временных рядов

Гибридные методы (алгоритм работы представлен на рисунке 1.4) прогнозирования временных рядов активно внедряются и используются во многих экономических секторах как в мире, так и в Российской Федерации, и сочетают в себе линейные и нелинейные модели прогнозирования и обработки больших

данных. К линейным моделям прогнозирования временных рядов и обработки больших данных относятся статистические модели, например, ARIMA [130] и GARCH [131]. В качестве нелинейных моделей прогнозирования временных рядов принято рассматривать интеллектуальные (нейросетевые) модели (например, модели MLN [139], DLN [140]), модели Грея [141] и ELM [142] (1.16). Данные модели широко применяются в прогнозировании временных рядов с возможностью обработки больших данных на региональном (например, в государственном комитете Республики Башкортостан по Чрезвычайным ситуациям) и локальном (в рамках предприятия) уровнях.

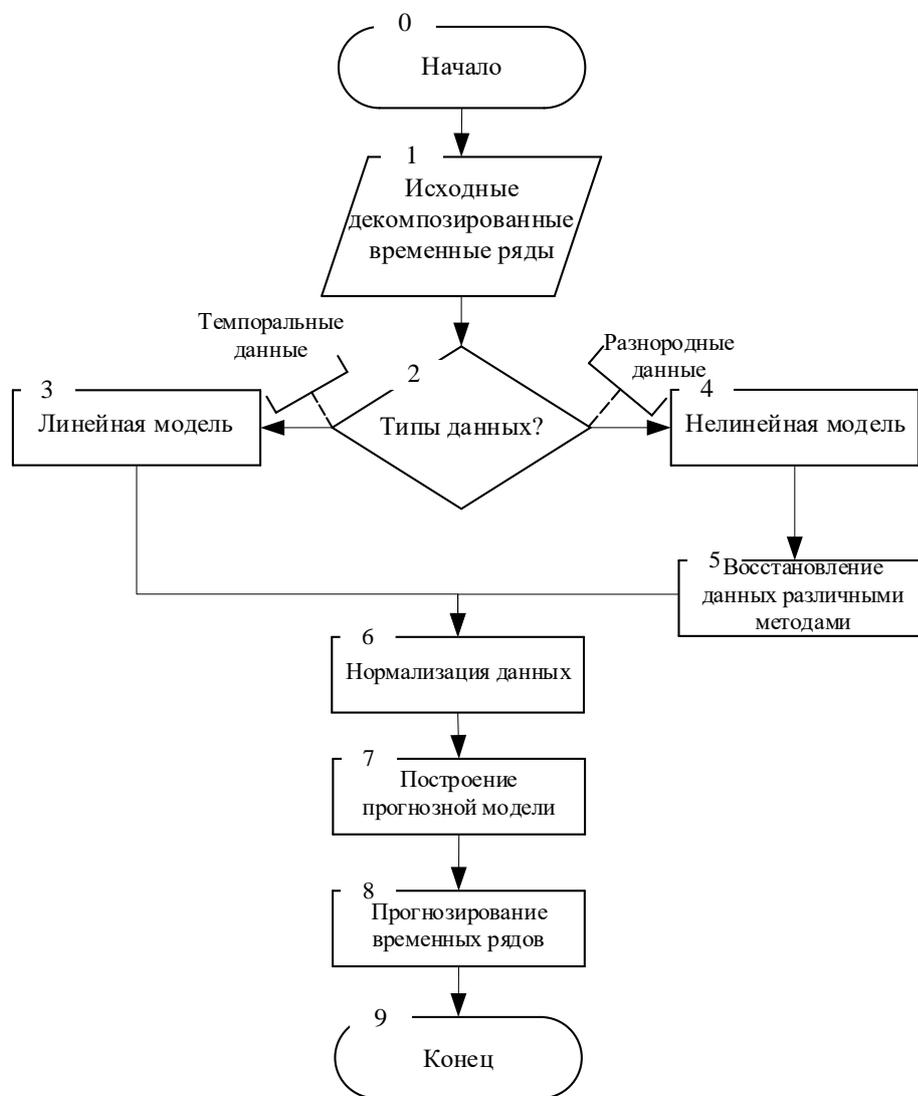


Рисунок 1.4 – Схема алгоритма предобработки данных и прогнозирование временных рядов на основе гибридного метода

Стандартный алгоритм предобработки данных и прогнозирования временных рядов на основе гибридного метода, представляемый в виде интеграции модели машинного и глубокого обучения, основанной на топологической структуре многослойного персептрона, имеет следующий вид:

$$Y = W_2\sigma(W_1x), \quad (1.16)$$

где W_1 – матрица весов между входом и скрытым слоем, σ – функция активации, W_2 – матрица весовых коэффициентов между скрытым слоем и выходом, x – вектор входного сигнала (входные данные, $x = [x_1, x_2, \dots, x_n]$). Суть такой интеграции заключается в повышении скорости подбора весовых коэффициентов синаптических связей за счет одного скрытого слоя нейронов, в котором веса между нейронами распределяются случайным образом и никогда не обновляются.

Эксперименты показывают, что подходы, используемые для декомпозиции временных рядов с исходной информацией и интеграции линейных и нелинейных моделей в процессе расчета прогнозных значений, являются важными элементами повышения точности методов прогнозирования и обработки больших структурированных данных, в том числе и прогнозирования временных рядов [129]. Например, в работе [141] для обработки больших данных и прогнозирования временных рядов использовалась дробно-дискретная модель Ферхюльста. Che L. и др. разработали гибридную модель для прогнозирования временных рядов и обработки больших данных на примере потребления электроэнергии [142]. Sulaiman S.M. и др. разработали новый гибридный метод прогнозирования временных рядов и обработки больших данных, основанный на эмпирической модовой декомпозиции (EMD) в тандеме с машиной экстремального обучения (ELM) на основе данных интеллектуальных счетчиков [143].

Большинство существующих исследований по разработке и применению гибридных методов прогнозирования временных рядов сосредоточены на повышении точности прогнозирования, но данные алгоритмы не умеют работать с большими разнородными данными: для них необходимо дописывать отдельную

предобработку данных, что не всегда является возможным. При этом в качестве существенного недостатка отмечается отсутствие достаточной универсальности и скорости прогнозирования.

1.2.4. Нейросетевые методы (модели) обработки больших данных с возможностью прогнозирования временных рядов

Нейросетевым методам обработки больших данных и прогнозирования временных рядов посвящено большое количество исследований, в которых подробно рассмотрены различные аспекты их применения [144–152]. Данные методы (рисунок 1.5) [147] прогнозирования временных рядов и обработки больших данных основываются на применении методов искусственного интеллекта – Deep Learning. Важно отметить, что по сравнению с ARIMA, MLR, моделями прогнозирования Грея и гибридными методами нейросетевые модели очень популярны, поскольку они хорошо справляются с нелинейностью и неоднородностью исходных данных. При этом модифицированные нейросетевые методы и модели, а также их практическая реализация и апробация приведены в главах 2 и 4 диссертационного исследования.

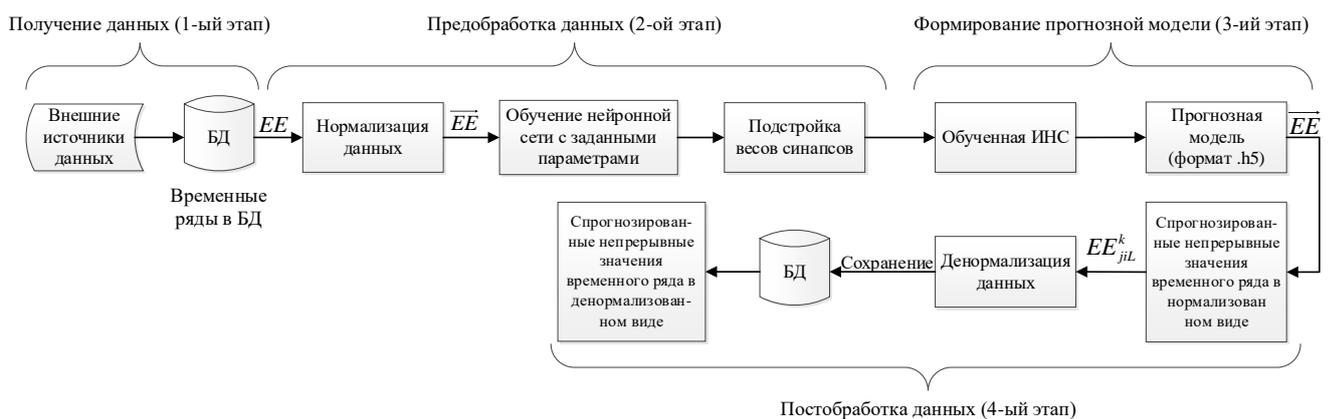


Рисунок 1.5 – Стандартная схема работы нейросетевых методов (моделей) прогнозирования и обработки больших данных

Например, в работах [148, 149] искусственные нейронные сети используются для прогнозирования временных рядов на основе обработки больших ретроспективных данных не только в компаниях в целом, но и в офисах в частности. В работе [150] рассматривается полупараметрическая нейронная сеть квантильной регрессии для прогнозирования временных рядов и обработки больших данных. В [151] предлагается метод долгосрочного прогнозирования временных рядов и обработки больших данных на основе системной динамики в рамках углеродно-нейтрального целевого показателя. В работе [152] использовали классификатор дрейфа для обработки большого объема сетевого потока данных в зависимости от типов сетей и прогнозирования временных рядов.

Основными преимуществами данных моделей являются скорость обучения и высокая точность прогнозирования временных рядов и обработки больших данных. Для обработки больших данных в целом, и прогнозирования временных рядов в частности, зачастую используется наиболее популярный метод обучения обратного распространения ошибки. Суть данного метода обучения сводится к распространению сигналов ошибки от выхода сети к ее входам с одновременным вычислением градиента. В качестве достоинства данного метода необходимо отметить устойчивость к аномалиям (выбросам) во временных рядах. При этом в классической вариации данного метода необходимо отметить следующие минусы: во-первых, заикленность сети, так как при резком возникновении аномалий во временном ряде сигмоидальная функция активации попадает в область насыщения, из-за чего производная величина приближается к 0, благодаря чему коррекция весовых коэффициентов синаптических связей сводится к минимуму, в конечном итоге, получаем неточные результаты после обучения нейронной сети; во-вторых, увеличивается вероятность попадания в локальные минимумы при малом шаге итерации, как следствие, прекращение обучения ИНС.

Тем не менее, к общим недостаткам этих традиционных и модифицированных моделей следует отнести отсутствие фиксированных характеристик, например, асимметричные циклы и случайные выбросы (аномалии)

наблюдений, существующие в реальных временных рядах, а также невозможность обработки разнородных данных в исходном виде (например, разнородный сетевой поток данных).

Исходя из обзора вышеперечисленных методов и моделей, можно сделать следующие выводы:

1. Статистические методы и модели обработки больших данных и прогнозирования временных рядов.

1.1. Основными *преимуществами* данных моделей и методов являются:

1.1.1. *Простота математической модели прогнозирования.* Это позволяет в кратчайшие сроки программно реализовать данную модель с целью краткосрочного прогнозирования временных рядов.

1.1.2. *Гибкость* (за исключением авторегрессионных методов и моделей из-за наличия большого количество прогнозируемых параметров). Если данные являются структурированными, то математическая модель является гибкой, что позволяет использовать ее во многих предметных областях.

1.2. Основные *недостатки*:

1.2.1. *Сложность определения вида функциональной зависимости.* Например, с помощью каких дополнительных параметров прогноз целевого параметра будет наиболее точным.

1.2.2. *Моделирование нелинейных процессов происходит с низкой точностью и на короткий промежуток времени*, что негативно сказывается на точности раннего прогнозирования из-за нелинейности временного ряда потребления электроэнергии.

1.2.3. *К каждой модели необходимо реализовывать отдельный модуль предобработки данных* (например, если данные являются разнородными и в них присутствует шум) с целью получения адекватных прогнозных значений в зависимости от предметной области.

2. Модели Грея обработки больших данных и прогнозирования временных рядов.

2.1. Основными *преимуществами* данных моделей являются:

2.1.1. *Простая реализация* из-за нетривиальности математического аппарата, что открывает возможность в кратчайшие сроки программно реализовать данную модель с целью прогнозирования временных рядов с малым прогнозным горизонтом.

2.1.2. *Высокая точность при невысоком горизонте прогнозирования* только в том случае, если данные полностью структурированы и найдена функциональная зависимость.

2.2. Основные *недостатки*:

2.2.1. При наличии разнородных данных зачастую наблюдаются *неточные результаты* при краткосрочном и раннем (заблаговременном) прогнозированиях из-за: недостаточно проработанной функция временного отклика TRF, основанная на гипотезе о фиксированном промежутке времени, что ограничивает адаптивность данных моделей к различным ситуациям; тривиальности своих исходных алгоритмов предобработки данных.

2.2.2. *Ресурсоемкость процесса расчета прогнозных значений* при большом потоке входных данных.

3. Гибридные методы обработки больших данных и прогнозирования временных рядов.

3.1. Основным *преимуществом* является прогнозная точность при краткосрочном прогнозировании из-за интеграции линейных и нелинейных моделей прогнозирования.

3.2. Основным *недостатками* является ресурсоемкость процесса прогнозирования, в том числе и процесса обучения и сложность реализации.

4. Нейросетевые модели и методы обработки больших данных и прогнозирования временных рядов.

Модели на нейронных сетях являются нелинейными, адаптивными и масштабируемыми. Также присутствует единообразие их анализа и проектирования.

4.1. Основные *преимущества*:

4.1.1. *Обрабатывать временные ряды* с нелинейными зависимостями, что существенно повышает прогнозную точность.

4.1.2. *Автоматическое определение целевого параметра и извлечение необходимых признаков из исходного датасета* без необходимости ручной настройки. Это повышает скорость обучения для прогнозирования временных рядов.

4.1.3. *Чем больше структурированных данных, тем выше точность.*

4.2. Основные *недостатки*:

4.2.1. *Асимметричные циклы и случайные выбросы* (аномалии) наблюдений, существующие в реальных временных рядах.

4.2.2. *Невозможность обработки разнородных данных в исходном виде* (например, разнородный сетевой поток данных).

4.2.3. *Сложность выбора архитектуры и реализации* модели в целом.

4.2.4. *Высокие требования к непротиворечивости обучающей выборки, сложность выбора алгоритма обучения и ресурсоемкость процесса их обучения.* Но для временных рядов, которые подвержены кризисным процессам лучшую эффективность дают нейросетевые модели.

Таким образом, рассмотренные преимущества и недостатки определили направление, цель и задачи диссертационного исследования.

1.3. Цель и задачи исследования

Все вышеперечисленные методы обработки больших разнородных и темпоральных данных и прогнозирования временных рядов подразделяются (таблица 1.3) на статистические (машинное обучение), модели Грея (машинное обучение), гибридные (машинное обучение) и нейросетевые (глубокое обучение).

Таблица 1.3 – Сравнительный анализ рассмотренных в главе 1 методов и моделей обработки больших данных и прогнозирования временных рядов

Параметр, по которому производится анализ	Статистические методы (машинное обучение)	Модели Грея (машинное обучение)	Гибридные методы (машинное обучение)	Нейросетевые методы (модели глубокого обучения)
Высокая точность при краткосрочном прогнозировании	+	+	+	+
Высокая точность при раннем (заблаговременном) прогнозировании	-	-	-	+
Гибкость под различные предметные области	+	-	-	-
Простота реализации	+	+	-	-
Прозрачность моделирования (доступность промежуточных вычислений для анализа)	-	-	-	+
Необходимость отдельной предобработки данных для каждой модели	+	+	+	+
Эффективность при прогнозировании на основе стационарного временного ряда	+	+	+	+
Эффективность при прогнозировании на основе нестационарного временного ряда	-	-	+	+
Эффективность при прогнозировании на основе временного ряда при наличии сезонной составляющей	-	-	+	+
Ресурсоемкость процесса обучения и прогнозирования	+	+	+	+
Скорость обучения – высокая/низкая (+/-)	+	+	-	-

Исходя из анализа, представленного в разделе 1.2 и таблице 1.3, у вышеперечисленных методов и моделей обработки больших данных и прогнозирования временных рядов имеются четыре серьезных общих ограничения в рамках рассматриваемой предметной области:

1) специфичность и узконаправленность: каждую модель и методы придется перепроектировать для каждого конкретного случая в рамках прогнозирования временных рядов;

2) ресурсоемкие модели прогнозирования;

3) нет формализованной и технической реализации обработки больших темпоральных и разнородных данных (например, данных, получаемых в различном виде с различных источников одновременно: счетчики, веб-ресурсы) с одновременным прогнозированием одномерных временных рядов на основе многомерных с учетом необходимых зависимостей на этапе предобработки данных, как следствие, необходимо выполнить разностную аппроксимацию входных параметров объекта (временного ряда);

4) отсутствие необходимого этапа предобработки информации, на выходе которого получаем качественный исходный датасет для обучения ИНС с учетом зависимостей между данными: информация о температуре с датчиков, выходные дни, конец месяца и т.д. Следовательно, отсутствуют сбор информации, ее организация и анализ, а также обнаружение и идентификация требуемого информационного блока в больших темпоральных массивах данных.

Предлагаемое в настоящем диссертационном исследовании решение учитывает и устраняет данные недостатки.

Таким образом, указанные проблемы и недостатки существующих решений определили цель диссертационного исследования, заключающаяся в повышении эффективности информационного сопровождения принятия управленческих решений в организации при проведении энергосберегающих мероприятий за счет нейросетевой обработки больших темпоральных данных с одновременным прогнозированием многомерных временных рядов.

Для достижения вышеуказанной цели, необходимо решить следующие задачи:

1. Разработать метод и модель обработки информации для формирования исходного датасета рекуррентной нейронной сети на основе правил фильтрации сетевого потока данных за счет обучения импульсной нейронной сети с целью прогнозирования значений потребления электроэнергии.

2. Разработать метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений.

3. Разработать структурную схему системы управления процессом поддержки принятия управленческих решений в электроэнергетике на основе предложенных методов и моделей.

4. Разработать прототип СППР на основе нейросетевой обработки больших темпоральных данных и прогнозирования временных рядов для предприятия электроэнергетической отрасли.

5. Оценить эффективность предложенных решений (обработка больших объемов сетевого потока данных и прогнозирование потребления электроэнергии) на примере предприятия электроэнергетической отрасли.

Выводы по главе 1

В результате проведенных исследований сделаны следующие выводы:

1. Проведенный анализ методов и моделей обработки больших данных и прогнозирования временных рядов в рамках темы исследования позволил сформулировать цель и задачи исследований, направленных на интеллектуальную (нейросетевую) обработку больших разнородных данных и прогнозирование временных рядов на примере обработки больших объемов сетевого потока данных и прогнозирования временных рядов (количества потребления электроэнергии).

2. Выявлено, что из-за огромного объема и присущей большим данным разнородности классические методы и модели, рассмотренные в пунктах 1.2.1–1.2.4, обработки данных и прогнозирования временных рядов не подходят для работы с ними по следующим причинам: специфичность и узконаправленность как методов, так и моделей; ресурсоемкость моделей прогнозирования; наличие разностной аппроксимации входных параметров объекта; отсутствие необходимо этапа предобработки данных.

3. Выполнена классификация методов обработки больших данных и прогнозирования временных рядов на статистические, модели Грея, гибридные и нейросетевые. Выявлено, что:

3.1. Основными недостатками статистических методов являются: невозможность корректного прогнозирования временных рядов при высоком значении горизонта прогнозирования; в случае предоставления на вход разнородных данных прогнозирование становится просто невозможным: нет четкой структурированности и фазсификации датасета, а для статистических методов это является основополагающим фактором обработки информации.

3.2. Основными недостатками моделей Грея являются неточные результаты при краткосрочном и раннем (заблаговременном) прогнозированиях из-за недостаточно проработанной функции временного отклика TRF (Time Response

Function), а также невозможность обработки больших разнородных данных в силу тривиальности своих исходных алгоритмов.

3.3. Основными недостатками гибридных методов являются низкая скорость прогнозирования временных рядов, отсутствие универсальности и невозможность обработки больших разнородных данных.

3.4. Основными недостатками нейросетевых моделей являются отсутствие фиксации характеристик (например, асимметричные циклы и случайные выбросы (аномалии) наблюдений), существующих в реальных временных рядах, а также невозможность обработки разнородных данных в исходном виде (например, разнородный сетевой поток данных).

4. Проанализированы и выявлены общие недостатки существующих методов обработки больших данных и прогнозирования временных рядов: нет универсальности (одновременного применения моделей и методов к различным процессам для принятия управленческих решений) относительно одновременной обработки больших данных и прогнозирования временных рядов нет возможности обработки больших разнородных данных в рамках исследуемой предметной области.

5. Исходя из литературного обзора, представленного в главе 1, и потребности реализации на практике (Стратегия развития Российского научного фонда до 2030 года, концепция «Цифровая трансформация 2030», программа «Развитие энергетики» до 2035 года и доктрина «Энергетическая безопасность Российской Федерации») выявлена необходимость в принятии качественных решений при управлении энергетическими организациями, например, при проведении энергосберегающих мероприятий, при расчете необходимых энерго мощностей для непрерывного обеспечения электроэнергией обслуживаемого района.

ГЛАВА 2. РАЗРАБОТКА МЕТОДОВ И МОДЕЛЕЙ ОБРАБОТКИ БОЛЬШИХ ТЕМПОРАЛЬНЫХ ДАННЫХ И ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Предложены методы интеллектуальной обработки больших разнородных и темпоральных данных и прогнозирования временных рядов с применением импульсной нейронной сети (третье поколение ИНС) и рекуррентной нейронной сети (второе поколение ИНС) на основе модифицированного метода обучения обратного распространения ошибки для информационной поддержки принятия решений на примере обработки больших объемов сетевого потока данных и прогнозирования временных рядов (количества потребляемой электроэнергии).

Предлагаемые методы и модели включают в себя:

– метод обработки сетевого потока данных на основе модели импульсной нейронной сети (третье поколение) для информационной поддержки принятия управленческих решений. При поступлении больших объемов сетевого потока данных через внешний сетевой интерфейс данные метод и модель позволяют качественно и эффективно осуществлять информационную поддержку принятия управленческих решений в электроэнергетике, в том числе и для формирования исходного датасета за счет формируемых на выходе правил фильтрации, низкой загруженности серверных ресурсов и скоростной обработки больших данных, что достигается методом обучения импульсной нейронной сети;

– метод прогнозирования временных рядов на основе модели рекуррентной нейронной сети второго поколения для принятия управленческих решений. Данный метод за счет модифицированного метода обучения и высокой точности прогнозных значений позволяет оперативно и эффективно осуществлять информационную поддержку принятия управленческих решений в электроэнергетике. Эффективность метода и модели экспериментально подтверждена на реальном массиве сетевого потока данных при этом средняя величина ошибки не превышала 7,8%.

2.1. Концептуальные положения решения задачи обработки и анализа больших темпоральных данных и прогнозирования временных рядов в электроэнергетике

Во-первых, исходя из анализа в главе 1, представленного в разделе 1.1, проблема интеллектуальной обработки больших разнородных и темпоральных данных и прогнозирования временных рядов в электроэнергетике имеет четко выраженные общие ограничения, связанные с предметной областью. Основное ограничение – затрудненный расчет точного прогнозного значения потребления электроэнергии. Вычислить данное значение возможно с достаточной степенью неопределенности, например, с помощью статистических или нейросетевых методов. При этом ошибку любой модели можно вычислить путем нахождения разности прогнозного и эталонного значения электропотребления, а также оценки возможных объемов энергосбережения (электросбережения) и условного разнесения полученных показателей по различным группам потребителей: первая – наиболее важные потребители электроснабжения; вторая – важные городские и производственные объекты; третья – частные и многоэтажные дома (п. 1.2.8 правила устройства электроустановок в седьмой редакции – ПУЭ-7). При этом наибольшего эффекта (наибольшей прогнозной точности) в рамках разрабатываемого прототипа СППР можно достичь с прогнозированием потребления электроэнергии в быту – вторая категория.

Во-вторых, в области электроэнергетики остаются значительные риски возникновения аварий на объектах. В частности, большое количество аварий происходит из-за оперативного персонала и перегрузки оборудования, а около половины из них связаны с ошибочными действиями лиц, принимающих решения. При учете последствий ошибок выделяют 3 этапа перехода от ошибок к авариям:

1) этап накопления. Нарушения и ошибки не проявляются внешне, предполагает незначительные отклонения сотрудниками от норм процесса;

2) этап инициирования аварии. Характеризуется наличием события, вызванного ранее совершенными ошибками, которые приводят, например, к избыточной нагрузке на оборудование;

3) этап взрыва. На данном этапе оператор не имеет средств и времени на предотвращение аварии.

В-третьих, в связи с повсеместной компьютеризацией и автоматизацией, в том числе и на предприятиях электроэнергетики, появляются несколько информационных систем, конкурирующих между собой за сетевой поток данных, в рамках которого передаются данные, присутствуют датчики, взаимодействующие сразу с несколькими информационными системами (передающие необходимые данные по расписанию, отдельно для разных информационных систем). Эти данные, курсирующие в рамках корпоративной сети предприятия, имеют большой процент зашумленности³, который необходимо отфильтровывать за счет правил фильтрации сетевого потока данных для формирования исходного датасета с целью обучения нейронных сетей. *При этом сформированный единожды датасет не может считаться постоянно оптимальным по причине изменения внешней среды: отказ в работе датчиков; появляются новые данные, а также новые системы, перенаправляющие информационные потоки, что приводит к увеличению задержки при передаче потоков данных; изменение количества датчиков.* Как следствие, необходимо анализировать сетевой поток данных на предмет необходимых параметров для прогнозирования потребления электроэнергии. Это позволит получать на выходе более качественный датасет, благодаря чему вырастет точность прогноза потребляемой электроэнергии, что позволит более эффективно использовать финансовые ресурсы энергетическим компаниям. Данный фактор может положительно повлиять не только на тарифные планы электроэнергии для

³ Информация по зашумленности основывается на проведенных экспериментах по фильтрации данных

физических и юридических лиц, но и предоставит возможность оптимизации бюджета энергетической компании.

Для реализации выше указанного предложен следующий алгоритм действий:

1. С помощью разработанного метода и модели обработки больших разнородных и темпоральных данных в виде сетевого потока данных с применением импульсной нейронной сети, обеспечить обработку получаемой информации через внешний сетевой интерфейс для формирования исходного датасета за счет формируемых на выходе ИмНС правил фильтрации, низкой загруженности серверных ресурсов и скоростной обработки больших данных, что достигается методом обучения (раздел 2.2) импульсной нейронной сети. На данном этапе также необходимо произвести согласование проводимых измерений по изменению видов и типов данных сетевого потока данных с установленными графиками передачи информации со счетчиков и датчиков (источниками данных). В случае выявления каких-либо отклонений, параметры которых приводятся в разделе 4.3, выполняется переобучение импульсной нейронной сети модифицированным методом обратного распространения ошибки, с заранее формирующейся архитектурой нейронной сети: общее количество нейронов и слоев, в том числе и промежуточных; количество нейронов на входном, промежуточном и выходном слоях.

Таким образом, реагирование на изменение вида и типа сетевого потока данных позволяет в кратчайшие сроки проанализировать сетевой поток данных и выработать специализированное правило его обработки (управляющего воздействия – фильтрации). Соответственно, после обнаружения неизвестного (прочего) типа аномалий сетевого потока данных происходит его фильтрация за счет выработки специализированных правил управляющего воздействия с последующим присвоением аномалии константы реагирования.

2. С помощью разработанного метода и модели прогнозирования временных рядов с применением рекуррентной нейронной сети, за счет модифицированного метода обучения и высокой точности прогнозных значений, осуществлять

информационную поддержку принятия управленческих решений в электроэнергетике. Данные метод и модель, позволяющие выполнять раннее прогнозирование, состоят из следующих этапов: принятие решения о прогнозировании (т.е. ЛПП принимает решение по количеству дней, на которые будет осуществлен прогноз); предобработка данных; формирование прогнозной модели; постобработка данных. Этап предобработки данных заключается в нормализации данных полученного датасета, сформированного на основе импульсной нейронной сети. При этом данные уже присутствуют в БД MySQL и являются структурированными за счет решения предыдущей задачи.

Общая функциональная структурная схема в рамках процесса автоматизации разработанного прототипа СППР приведена в главе 3 и включает в себя:

- обеспечение автоматического взаимодействия между каждым модулем и веб-частью;
- реализация платформы для бесконфликтной и взаимосвязанной работы модулей между собой;
- создание связей между базой данных MySQL и модулями для дальнейшего хранения передаваемой информации;
- обработка и фильтрация больших объемов сетевого потока данных, а также прогнозирование временных рядов для поддержки принятия решений в области электроэнергетики.

Учитывая, что управленческие решения могут быть связаны с выполнением различных по степени важности мероприятий, например, резервированием дополнительных мощностей электроэнергетического предприятия на основе прогнозных значений потребления электроэнергии, или внедрением правил фильтрации сетевого потока данных для решения задач оптимизации, предлагается четырехкаскадный контур управления в виде четырех замкнутых контуров управления.

Таким образом, за счет разработанных методов и моделей стало возможным раннее и точное (с минимальным процентом отклонений, глава 4) прогнозирование потребления электроэнергии, что позволяет повысить точность экспертных заключений (глава 3) и уменьшает роль человеческого фактора при принятии решений в различных ситуациях (глава 3).

2.2. Разработка метода и модели обработки сетевого потока данных с применением импульсной нейронной сети

Предлагаемые метод и модель обработки больших темпоральных данных на примере больших объемов внешнего сетевого потока данных, представленного в виде привязанной к дате и времени информации, передаваемой с датчиков и счетчиков электрических подстанций, основаны на применении ИмНС третьего поколения. Следовательно, модель обработки больших темпоральных данных – это есть импульсная нейронная сеть. Необходимо отметить, что приоритет по обработке больших темпоральных данных отдается импульсной нейронной сети из-за ее высокой скорости работы при обработке данных, динамичности и многозадачности.

Из-за необходимости повышения качества принимаемых решений в электроэнергетике, в том числе и на основе интеграции информационно-вычислительных и искусственных нейронных сетей и прогнозных значений в целом, одним из основных параметров воздействия на принятие решений в электроэнергетике являются данные (информация, поступающая по сетевым каналам через внешний сетевой интерфейс), передаваемые от хоста (h) к клиенту (c), в нашем случае – на физические серверы кластера за определенный период времени t . В этой связи введем обозначения $N_c^h(t)$ – поступающие темпоральные

данные⁴ (рисунок 2.1) с внешнего сетевого интерфейса, передаваемые с хоста h (счетчиков и датчиков подстанции) клиенту c за определенный период времени t , при этом $h = 1, \dots, n$ и $c = 1, \dots, k$ (n – общее количество хостов, с которых передается сетевой поток данных, k – количество физических серверов в кластере (рисунок 2.2) – клиентов).

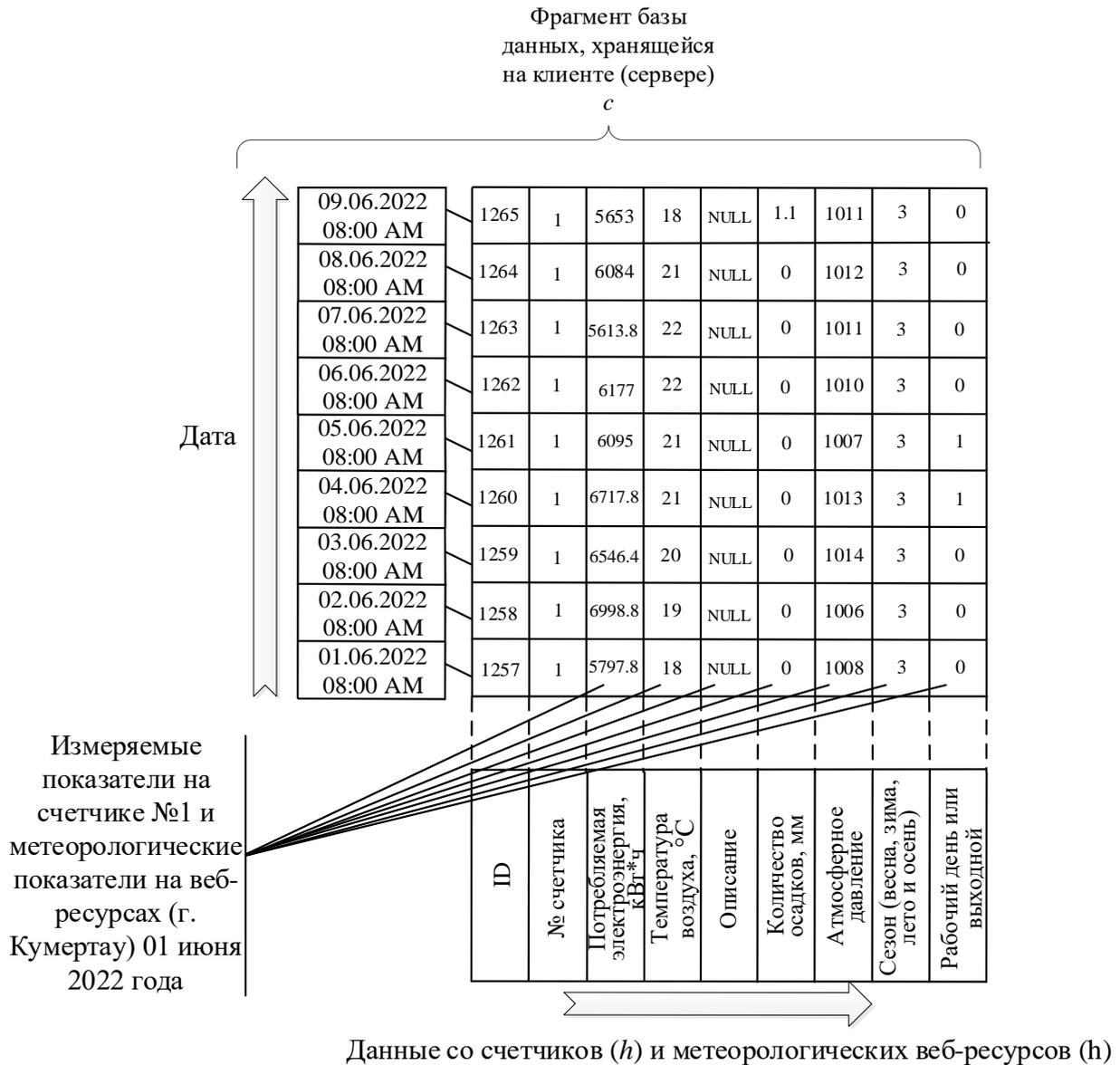


Рисунок 2.1 – Фрагмент данных $N_c^h(t)$, отправляемых на вход импульсной нейронной сети

⁴ Дата измерения, № счетчика, потребляемая электроэнергия, температура воздуха, количество осадков, выходной день или нет, атмосферное давление и сезон, доля от общих данных – 90%

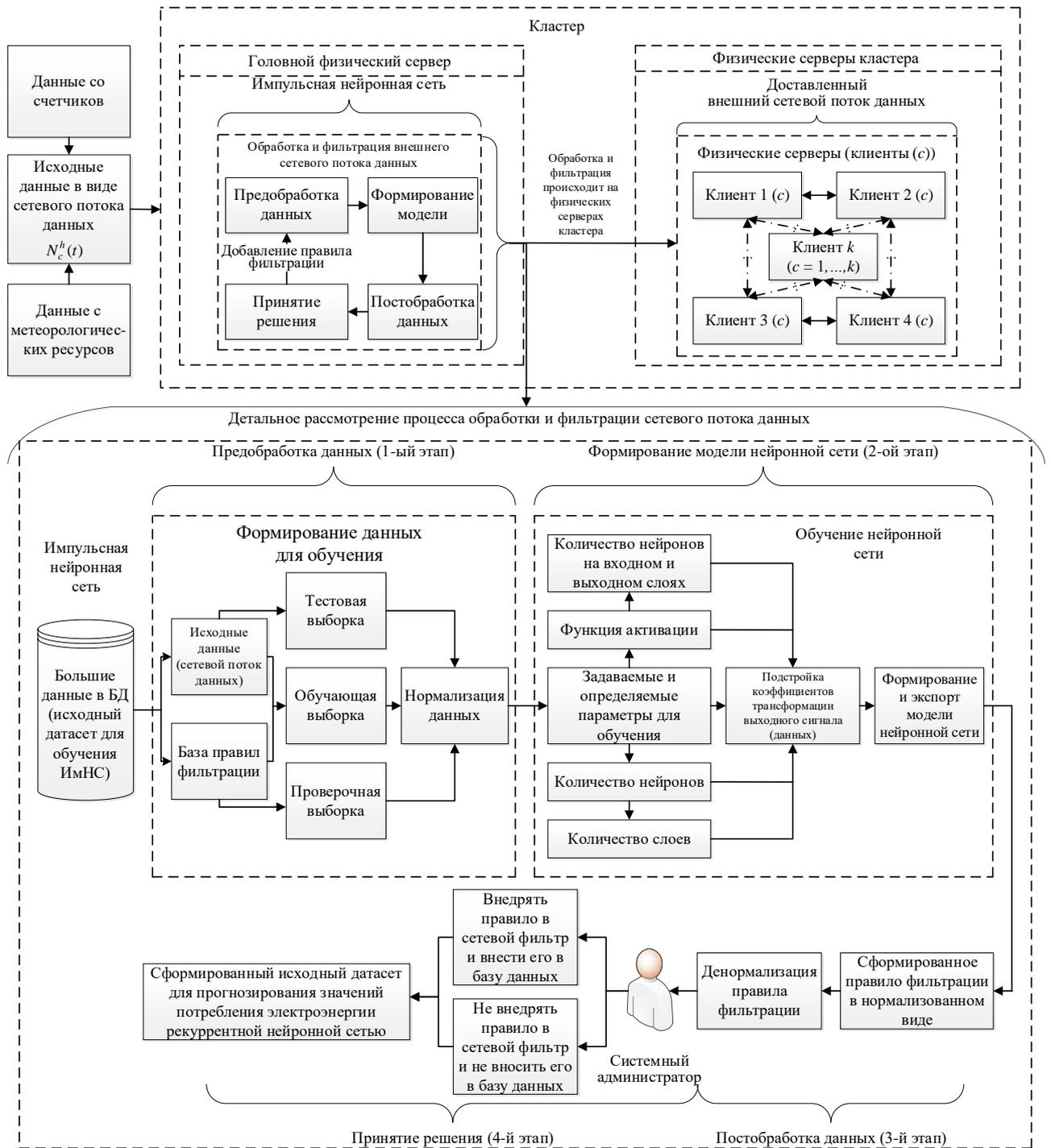


Рисунок 2.2 – Структура кластера (новая сеть доставки контента) и схема формирования базы правил для формирования исходного датасета

Исходя из введенных ранее обозначений, задача обработки больших темпоральных данных заключается в том, чтобы за определенный период времени t обработать максимальное количество данных $N_c^h(t)$ с внешнего сетевого

интерфейса для формирования исходного датасета за счет правил фильтрации и минимизации рисков несвоевременного реагирования на изменение ситуации, в частности, потери работоспособности клиента c . Решать данную задачу предлагается в четыре этапа, что представлено на рисунке 2.2 (блоки импульсной нейронной сети).

На первом этапе, выполняющемся каждый час – i (согласно проведенным исследованиям и по причине установленной периодичности передачи данных [153], а также изменения видов и типов входящего сетевого потока данных), т.е. $i = 24$ обучения в сутки, мы формируем архитектуру нейронной сети, осуществляя выбор параметров импульсной нейронной сети согласно формулам (2.1–2.5): наработанная база правил фильтрации; общее количество нейронов ($NINN_{total}$) и слоев (NNL_{total}), в том числе и промежуточных (NNL_h); количество нейронов на входном ($NINN^{il}$), промежуточном ($NINN^{hl}$) и выходном слоях ($NINN^{ol}$).

Далее рассматриваются вопросы расчета общего количества нейронов и слоев в импульсной нейронной сети, количества нейронов на входном и выходном слоях, формирования структуры (модели) нейрона ИмНС. Для выработки правил фильтрации и определения их количества описывается и реализуется обучение ИмНС модифицированным методом обратного распространения ошибки.

Общее количество нейронов и слоев в импульсной нейронной сети определяется по разработанным формулам (2.1–2.5):

$$R_c^{sc} = \begin{cases} R_{ram} = R - R1 \\ R_{cpu} = C - C1 \end{cases} \quad (2.1)$$

где R_c^{sc} – количество свободных ресурсов на физическом сервере (c) в кластере (sc) перед обучением нейронной сети, R_{ram} – количество килобайт свободной оперативной памяти каждого физического сервера в кластере перед обучением нейронной сети, R – суммарная в килобайтах оперативная память на физическом сервере в кластере перед обучением нейронной сети, $R1$ – используемая в килобайтах оперативная память на физическом сервере в кластере перед обучением

нейронной сети, R_{cpu} – количество свободных ресурсов (в %) CPU физического сервера в кластере перед обучением нейронной сети, C – суммарные (в %) ресурсы CPU физического сервера в кластере перед обучением нейронной сети, $C1$ – используемые в % ресурсы CPU физического сервера в кластере перед обучением нейронной сети. Обозначим R_m размер памяти, выделяемый под каждый нейрон, тогда общее число нейронов может быть вычислено по формуле $NINN_{total} = \frac{R_{ram}}{R_m}$.

В результате исследований установлено, что целесообразно под каждый нейрон выделить 100 кб RAM и установить ограничение R_{cpu} не более 1% нагрузки на CPU на каждом физическом сервере в кластере, исходя из чего общее число нейронов $NINN_{total}$ определяется соотношением

$$NINN_{total} = \frac{R_{ram}}{9.537 \cdot 10^{-5}}. \quad (2.2)$$

Таким образом, общее число слоев NNL_{total} определяется по соотношению

$$>NNL_{total} =>NNL_i +>NNL_h +>NNL_o, \quad (2.3)$$

где NNL_i – входной слой нейронной сети, NNL_h – промежуточные слои нейронной сети (2.4–2.5), NNL_o – выходной слой нейронной сети. Тогда

$$NINN^{hl} = NINN_{total} - (NINN^{il} + NINN^{ol}), \quad (2.4)$$

где $NINN^{hl}$ – количество нейронов скрытых промежуточных слоев нейронной сети, $NINN^{il}$ – количество нейронов входного слоя нейронной сети, $NINN^{ol}$ – количество нейронов выходного слоя нейронной сети. Исходя из найденного значения $NINN^{hl}$ вычисляются NNL_h :

$$>NNL_h = \frac{NINN^{hl}}{NINN^{il}}. \quad (2.5)$$

Для нумерации нейронов в описанных выше переменных $NINN^{hl}$, $NINN^{il}$ и $NINN^{ol}$ введем индекс k – порядковый номер нейрона в слое.

Количество нейронов на входном и выходном слоях. Количество нейронов на входном слое напрямую зависит от количества подаваемых данных (в нашем случае – от поступающего на внешний сетевой интерфейс потока данных в виде сетевых пакетов) и определяется следующим образом:

$$NINN^{il} = \psi(N_{np}), \quad (2.6)$$

где N_{np} – количество входящих сетевых пакетов в секунду, ψ – функция зависимости количества нейронов от количества подаваемых данных: один нейрон = один сетевой пакет с темпоральными данными.

В случае с входным и выходным слоями функция ψ представлена следующим правилом: число нейронов зависит от количества сетевых правил для фильтрации сетевого потока данных (соответственно, одно правило = один нейрон).

Архитектура разработанной импульсной нейронной сети для обработки больших разнородных и темпоральных данных представлена на рисунке 2.2.

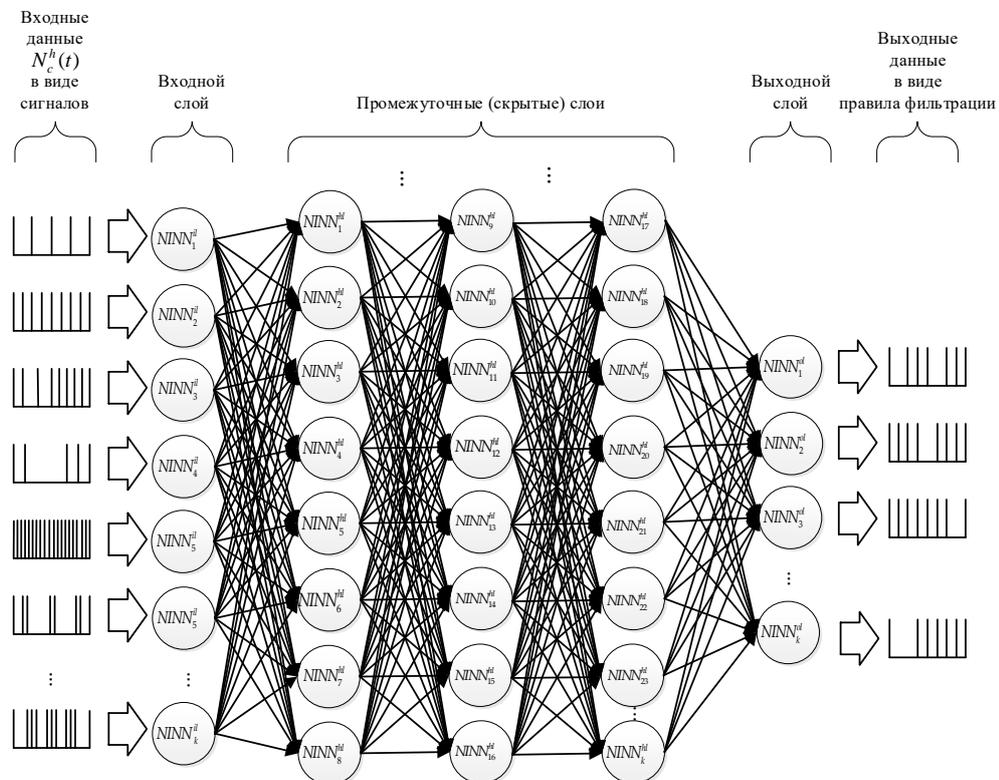


Рисунок 2.3 – Пример структуры разработанной ИМНС

Изначально (рисунок 2.3) данные (рисунок 2.1), представленные в виде сетевого потока данных и база правил фильтрации (рисунок 2.2) для последующей генерации соответствующего правила на выходе ИмНС, считываются с внешнего сетевого интерфейса каждого физического сервера с последующим преобразованием (кодированием) сетевых пакетов для их последующего представления в ИмНС методом порядка следования импульсов (аналог нормализации данных (2.7)), что представлено на рисунке 2.4: от расположения того или иного спайка (после преобразования сетевых пакетов) зависит скорость передачи данных и их обработка нейронной сетью. Но в связи с тем, что нейронные сети работают со значениями от 0 до 1 более корректно, т.е. на выходе получаем более точный результат, было принято решение реализовать специализированный кодировщик данных: числовые (таблица 2.1, формула (2.7)) и символьные (в случае IPv6 и правил фильтрации, таблицы 2.1–2.2) значения преобразовываются в диапазон от 0 до 1, что представлено в виде импульсов (данных).

$$LN_c^h(1) = \frac{LN_c^h(0) - N_{\min}}{N_{\max} - N_{\min}}, \quad (2.7)$$

где $LN_c^h(0)$ – текущее значение длины данных $N_c^h(t)$ в структуре сетевого пакета, $LN_c^h(1)$ – нормализованное значение длины данных $N_c^h(t)$ в структуре сетевого пакета, N_{\min} и N_{\max} – минимальное и максимальное значение длины данных в структуре сетевого пакета по всем данным, подаваемым на входной слой нейронной сети для всех $N_c^h(t)$.

Денормализуем данные по обратной формуле нормализации:

$$LN_c^h(0) = LN_c^h(1) \cdot (N_{\max} - N_{\min}) + N_{\min}. \quad (2.8)$$

Таблица 2.1 – Заданные значения для кодировщика с целью дальнейшего представления данных в ИмНС

Значение поля сетевого пакета	Ключевое слово (протокол)	Значение для кодировщика
0	Reserved	0

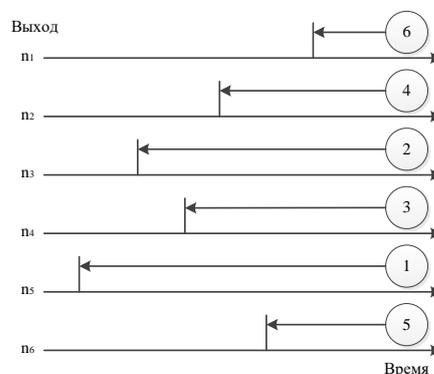
Таблица 2.1 – Окончание

Значение поля сетевого пакета	Ключевое слово (протокол)	Значение для кодировщика
1	ICMP	0,1
2	IGMP	0,2
4	IP	0,3
6	TCP	0,4
17	UDP	0,5
89	OSPF	0,6

Таблица 2.2 – Заданные буквенные значения (в случае с IPv6) для кодировщика с целью дальнейшего представления данных в ИмНС

Значение поля сетевого пакета	Значение для кодировщика
a	0,01
b	0,02
c	0,03
d	0,04
e	0,05

После формирования структуры нейронной сети и кодирования информации необходимо обучить ИНС. Запуск процесса обучения ИмНС осуществляется (в нашем случае) планировщиком «Cron» [154] с целью выполнения следующих задач: анализ сетевого потока данных, его распознавание и фильтрация с последующей записью полученных правил фильтрации сетевого потока данных в БД. В результате обучения ИмНС (более детальное описание представлено в пункте 2.2.2) получаем преобразованное правило фильтрации для формирования исходного датасета.

Рисунок 2.4 – Пример реализации метода порядкового следования импульсов (т.е. данных – $N_c^h(t)$) в разрабатываемой ИмНС

На рисунке 2.4 n_1, n_2, \dots, n_6 – выходы нейрона первого слоя, при этом 1, 2, ..., 6 – номера выходов.

Далее рассматриваются архитектура разработанной модели нейрона в ИмНС (пункт 2.2.1) и ее метод обучения (пункт 2.2.2).

2.2.1. Архитектура модели нейрона ИмНС

В импульсной нейронной сети для обработки больших разнородных и темпоральных данных за основу берется модель Фитц Хью-Нагумо [155–157], которая представляет собой упрощенную модель, воспроизводящую основные свойства волн возбуждения в модели Ходжкина-Хаксли [158] и содержит две переменные: быструю переменную, соответствующую мембранному потенциалу в полной модели, и медленную переменную, заключающуюся в значении восстановления мембранного потенциала. Внесены изменения, суть которых состоит в том, что в отличие от оригинальной модели добавлена возможность реагирования каждого нейрона на внешние воздействия (в нашем случае — на изменение видов и типов сетевого потока данных), как следствие, на динамику модели (архитектура нейрона представлена на рисунке 2.5) с нелинейным поведением восстанавливающей переменной.

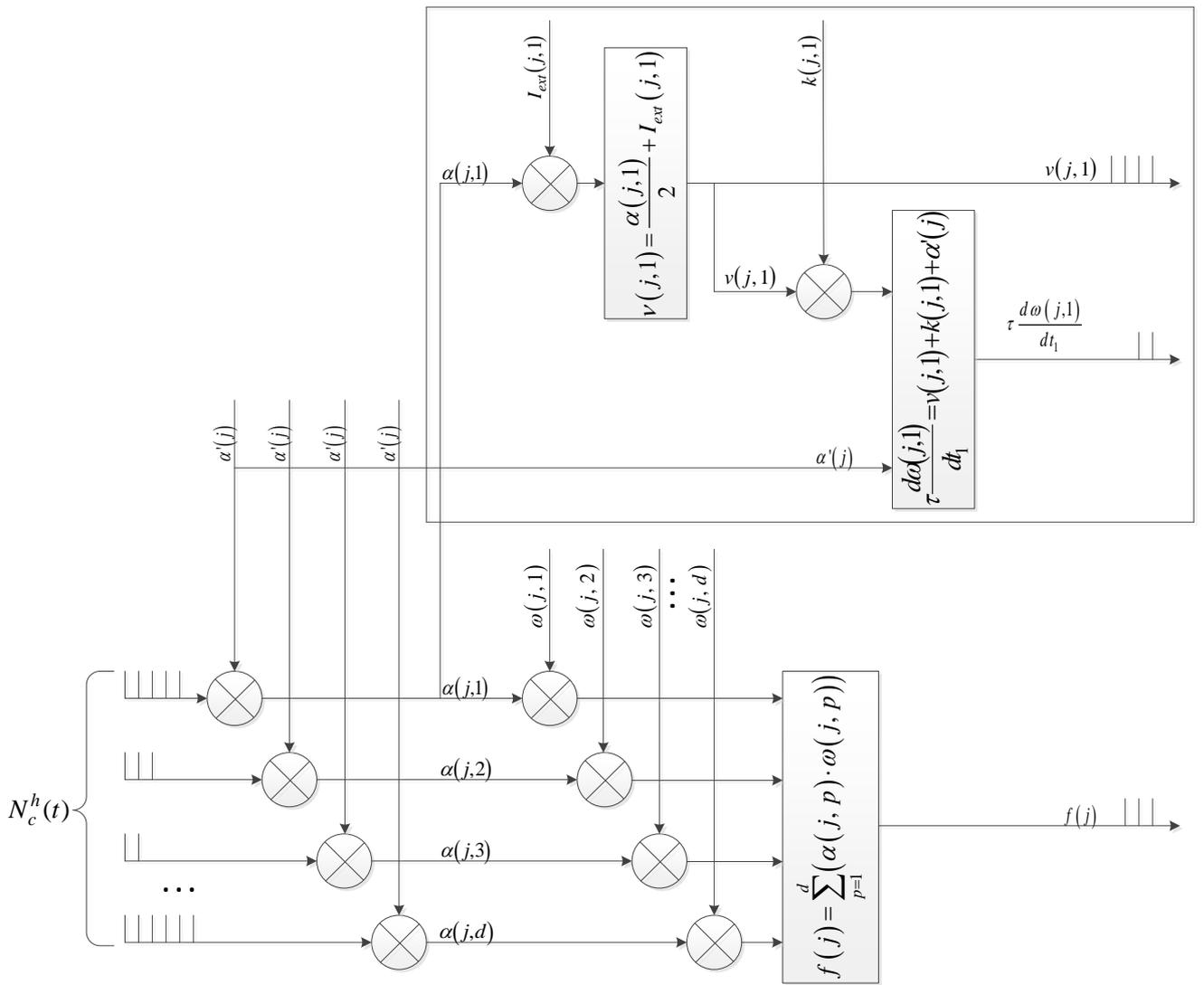


Рисунок 2.5 – Нейрон, имеющий многослойную структуру узла ИМНС

В модели нейрона (рисунок 2.5): j – порядковый номер нейрона; $j = (\overline{1, \dots, m})$, где m – это общее число нейронов; $\alpha'(j)$ – начальный потенциал j -го нейрона; p – порядковый номер импульса, $p = (\overline{1, \dots, d})$, где d – это количество импульсов; $\alpha(j, p)$ – потенциал j -го нейрона во время угасания p -го импульса; $\omega(j, p)$ – постоянная константа восстановления потенциала j -го нейрона в результате p -го импульса; t_p – время угасания p -го импульса; $f(j)$ – значение функции активации j -го нейрона:

$$f(j) = \sum_{p=1}^d (\alpha(j, p) \cdot \omega(j, p)) \quad (2.9)$$

Данная модель нейрона описывается следующей системой уравнений:

$$\begin{cases} v(j, p) = \frac{\alpha(j, p)}{2} + I_{ext}(j, p) \\ \tau \frac{d\omega(j, p)}{dt_p} = v(j, p) + k(j, p) + \alpha'(j) \end{cases}, \text{ при } p = (\overline{1, \dots, d}), j = (\overline{1, \dots, m}), \quad (2.10)$$

где $I_{ext}(j, p)$ – постоянная константа внешнего воздействия p -го импульса на j -й нейрон; t_p – время угасания p -го импульса; $k(j, p)$ – постоянная константа реагирования j -го нейрона на p -й импульс; $v(j, p)$ – функция динамики мембранного потенциала j -го нейрона от воздействия p -го импульса; τ – временная постоянная корреляции шума.

Таким образом, реагирование на изменение вида и типа сетевого потока данных (таблица 2.3) позволяет в кратчайшие сроки проанализировать сетевой поток данных и выработать специализированное правило его обработки (управляющего воздействия – фильтрации). Соответственно, после обнаружения неизвестного типа (аномалий) сетевого потока данных [159] происходит его фильтрация за счет выработки специализированных правил управляющего воздействия с последующим присвоением аномалии константы реагирования.

Таблица 2.3 – Значения константы реагирования $k(j, p)$

Тип аномалий в сетевом потоке данных	Коэффициент
UDP-флуд	0,001
DNS-усиление	0,002
HTTP-флуд	0,003
Все виды ICMP-флуда	0,004-0,120
MAC-флуд	0,121
SYN-флуд	0,122
NTP-усиление	0,123
TCP-флуд (Reset)	0,124
Source-флуд	0,125

Тип аномалий в сетевом потоке данных	Коэффициент
VoIP-флуд	0,126
Прочие типы аномалий сетевого потока данных	0,127-0,305

Для выработки правил фильтрации необходимо обучить импульсную нейронную сеть, что представлено в пункте 2.2.2.

2.2.2. Обучение ИмНС модифицированным методом обратного распространения ошибки

Наиболее часто в импульсных нейронных сетях используют метод обучения STDP (модифицированное правило Хебба) по причине простоты математической и программной реализации, а также достаточно большой эффективности при решении задач классификации и кластеризации. Но из-за использования в данном методе исключительно локальных правил обучения, основным недостатком является сложность осуществления обратного прохода при обучении на сигнале ошибки. Соответственно, ошибочный сигнал доступен строго на выходном слое импульсной нейронной сети и поток данных, передаваемый внутри нейронной сети, остается однонаправленным. В связи с вышеизложенным предлагается метод обратного распространения ошибки для импульсных нейронных сетей. Основной проблемой является бинарный характер входов и выходов импульсных нейронов, что не дает возможности использовать градиентные методы обучения. Основная суть предлагаемого в диссертации метода обучения заключается в накоплении потенциала в мембране нейрона за счет полученных на вход импульсов (спайков). Исходя из этого, если накопленный потенциал достигает порогового значения (в нашем случае – 1), то нейрон активируется и пропускает импульс (спайк) в следующий слой с последующим уменьшением своего потенциала (обозначим

данный процесс как $P_{j,p}$) до минимального значения (в нашем случае – 0), что описывается следующими уравнениями:

$$P_{j,p} = \frac{d\alpha(j,p)}{dt} = -\alpha(j,p) + I(j,p), \quad (2.11)$$

$$I(j,p) = w_{jp} \cdot \sum_{i=1}^d \varphi(p,i), \quad (2.12)$$

$$\varphi(p,i) = \begin{cases} 1, & \text{если } t_p = t_i \\ 0, & \text{иначе} \end{cases}, \quad (2.13)$$

в которых: $I(j,p)$ – взвешенная сумма входных импульсов (спайков) в текущий момент времени (t_p) на каждом на j -ом нейроне; $\varphi(p,i)$ – индикатор (переключатель) уровня потенциала в процессе $P_{j,p}$; w_{jp} – весовой синаптический коэффициент на каждом на j -ом нейроне.

Важно отметить, что группы импульсов представляют собой входные паттерны. Соответственно, чтобы сгенерировать импульсные входы, необходимо их распределить в специализированные спайковые группы по Пуассону [160–162] и подать в сеть. Исходя из 2.12 и учитывая, что $I(j,p)$ – взвешенная сумма входных импульсов (спайков) в текущий момент времени (t_p) на каждом на j -ом нейроне, общая взвешенная сумма входных спайковых групп S описывается следующим образом:

$$S = \sum_{j=1}^{m-1} \sum_{p=1}^d I(j,p). \quad (2.14)$$

При этом утечка потенциала на выходном слое учитывается разницей в утечке потенциала в текущий момент времени (t). Исходя из вышеописанного получаем функцию активации $f(j,p)$:

$$f(j,p) = \sum_{i=1}^d \exp\left(-\frac{t_p - t_i}{P_{j,p}}\right), j = \overline{(1, \dots, m)}, p = \overline{(1, \dots, d)}. \quad (2.15)$$

Сами весовые коэффициенты w_j рассчитываются по модифицированному методу обратного распространения ошибки [163–169].

Таким образом, за счет имеющегося потенциала у нейронов, их меньшего количества (в отличие от ИНС других поколений) и разработанного метода порядка следования импульсов, благодаря которому отправляется не непрерывная величина весового коэффициента, а 1 бит от весового коэффициента, предложенный метод обучения существенно увеличивает скорость и точность обучения импульсной нейронной сети для обработки больших разнородных и темпоральных данных с целью поддержки принятия решений на примере обработки больших объемов сетевого потока данных, что будет экспериментально показано в главе 4.

На основе данного метода обучения предлагается разработать математическую модель обучения искусственной нейронной сети второго поколения (раздел 2.3) и модифицировать ее с целью повышения точности прогнозирования временных рядов.

2.3. Разработка метода и модели прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети

В связи с тем, что прогнозирование является одной из неотъемлемых частей в рамках стратегии развития современных предприятий по многим причинам, в частности, по предотвращению аварийных ситуаций или экономии бюджета, возникает необходимость не только в обработке больших темпоральных данных, но и в одновременном расчете прогнозных значений в той или иной области. В рамках данной диссертационной работы прогнозирование осуществлялось на примере потребления электроэнергии энергетических предприятий Республики Башкортостан с учетом обработки больших темпоральных данных за счет модели модифицированной рекуррентной нейронной сети. Согласно экспериментальным исследованиям [170–175], рекуррентная нейронная сеть наиболее точно подходит для решения регрессионных задач за счет своего адаптивного математического аппарата.

Введем обозначения: EE_{ji}^k – значение потребляемой электроэнергии, измеренное на k -ом электросчетчике на i -ую дату j -го года. Здесь $k = 1, \dots, n$, где n –

количество электросчетчиков, участвующих в расчетах, j – номер года, i – конкретная дата измерения.

Задача раннего (заблаговременного) прогнозирования заключается в том, чтобы в конкретный текущий i -ый день измерения вычислить прогнозное значение потребляемой электроэнергии на $i+1$ день, то есть EE_{ji+1}^k , а также на $i+2$ (EE_{ji+2}^k), $i+3$ (EE_{ji+3}^k), $i+4$ (EE_{ji+4}^k), ..., $i+l$ (EE_{ji+l}^k), для любого $k = 1, \dots, n$.

Сам процесс прогнозирования, в который входит и обработка больших темпоральных данных, предлагается проводить в четыре этапа (рисунок 2.6):

1. Принятие решения о прогнозировании. Т.е. ЛПР принимает решение по количеству дней, на которые будет осуществлен прогноз.
2. Предобработка данных.
3. Формирование прогнозной модели.
4. Постобработка данных.

Частично разрабатываемые модель и метод применялись в рамках ранее проведенных исследований [163, 171]. Результаты данных исследований также подтверждают эффективность предложенного метода и модели прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений.

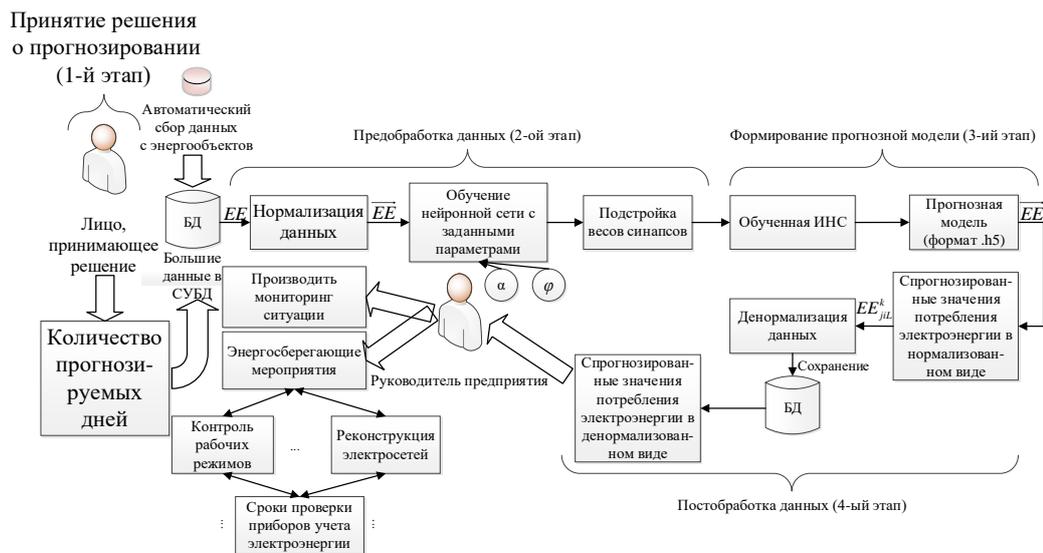


Рисунок 2.6 – Схема применения нейронной сети для прогнозирования значений потребления электроэнергии

На первом этапе ЛПР определяет на какое количество дней рассчитывать прогнозные значения потребления электроэнергии (в зависимости от месяца и различных задач). На втором этапе, на основе ранее описанных правил и формул из пункта 2.2.2, производится определение значений параметров архитектуры нейронной сети, а также обучение ИНС:

1. **Общее количество (2.16–2.20) нейронов N_{total} и слоев L_{total} .** Для обучения ИНС необходимы следующие ресурсы физического сервера: центральное процессорное устройство (центральный процессор, CPU), запоминающее устройство с произвольным доступом (оперативная память, RAM), твердотельный накопитель (SSD) или жесткий диск (HDD). Первоначально вычисляется количество свободных ресурсов на каждом физическом сервере кластера:

$$R_{free} = \begin{cases} R_{RAM} = R_{RAMsumm} - R_{RAMusage} \\ R_{CPU} = R_{CPUsumm} - R_{CPUusage} \end{cases} \quad (2.16)$$

В формуле (2.16):

- R_{free} – количество свободных ресурсов на физическом сервере (c) перед обучением нейронной сети;
- R_{RAM} – количество килобайт свободной оперативной памяти физического сервера перед обучением нейронной сети;
- $R_{RAMsumm}$ – суммарная в килобайтах оперативная память на физическом сервере перед обучением нейронной сети;
- $R_{RAMusage}$ – используемая в килобайтах оперативная память на физическом сервере перед обучением нейронной сети;
- R_{CPU} – количество свободных ресурсов (в %) CPU физического сервера перед обучением нейронной сети;
- $R_{CPUsumm}$ – суммарные ресурсы (в %) CPU физического сервера перед обучением нейронной сети;

– $R_{CPUusage}$ – используемые ресурсы в % CPU физического сервера перед обучением нейронной сети.

Аналогично импульсной нейронной сети (раздел 2.1), под каждый нейрон выделено 100 кб RAM и установлено ограничение в виде 1% нагрузки на CPU каждого физического сервера в кластере, исходя из чего N_{total} определяется соотношением (2.17) при использовании R_{RAM} :

$$N_{total} = \frac{R_{RAM}}{100}. \quad (2.17)$$

Таким образом, L_{total} определяется по соотношению

$$L_{total} = L_i^{in} + L_h + L_o^{out}, \quad (2.18)$$

где L_i^{in} – входной слой нейронной сети (равен 1), L_h – промежуточные слои нейронной сети (2.19–2.20), L_o^{out} – выходной слой нейронной сети (равен 1). При этом $i = 1$, $h = 1, \dots, tnil$ ($tnil$ – суммарное количество промежуточных слоев в нейронной сети), $o = 1$.

2. Количество нейронов во входном слое зависит от глубины прогнозирования значений временного ряда. Данное значение подбирается экспериментальным путем.

3. Количество нейронов в выходном слое напрямую зависит от значения горизонта прогнозирования ($EE_{ji+1, \dots, 31, L}^k$).

4. Количество (N_{hl}) нейронов (EE_{jk}) в промежуточных (скрытых) слоях нейронной сети:

$$N_{hl} = N_{total} - (N_{il} + N_{ol}), \quad (2.19)$$

где N_{il} – количество нейронов входного слоя нейронной сети, N_{ol} – количество нейронов выходного слоя нейронной сети.

5. Количество промежуточных (скрытых) слоев $tnil$, вычисляется исходя из найденного значения N_{hl} :

$$tnil = \frac{N_{hl}}{N_{il}}. \quad (2.20)$$

6. **Функция активации на промежуточных слоях** является сигмоидальной. Основной причиной использования данной функции является дифференциация на всей оси абсцисс [163], что повышает точность выходных значений нейрона. При этом на выходном слое используется линейная функция активации по причине повышения точности прогнозных значений на каждом нейроне выходного слоя [164].

7. **Количество данных по выборкам:**

а) Обучающая выборка: (train, 60%). Размерность обучающей выборки $Set_{training}$ (2.21) зависит от двух вышеобозначенных параметров: глубины прогнозирования и количества признаков (параметров) в исходных данных (таблица 4.14, глава 4):

$$Set_{training} = \begin{pmatrix} D^k(1) & k(1) & EE_{ji}^k(1) & AT_{ji}^k(1) & Pr_{ji}^k(1) & AP_{ji}^k(1) & Season_{ji}^k(1) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ D^k(m) & k(p_m) & EE_{ji}^k(en_m) & AT_{ji}^k(C_m) & Pr_{ji}^k(P_m) & AP_{ji}^k(A_m) & Season_{ji}^k(G_m) \end{pmatrix}, \quad (2.21)$$

где: $D^k(m)$ – дата измерения k -го счетчика (индекс), с учетом выходных и праздничных дней; m – количество (индекс) дат; $k(p_m)$ – номер счетчика; p_m – количество счетчиков; en_m – порядковый номер измерения значения потребления электроэнергии; AT_{ji}^k – температура воздуха k -го счетчика на i -ую дату j -го года; C_m – порядковый номер измерения температуры воздуха; Pr_{ji}^k – количество осадков на местности k -го счетчика на i -ую дату j -го года; P_m – порядковый номер измерения количества осадков; Ar_{ji}^k – атмосферное давление на местности k -го счетчика на i -ую дату j -го года; A_m – порядковый номер измерения атмосферного давления; $Season_{ji}^k$ – сезон года (зима, весна, лето или осень) на местности k -го счетчика на i -ую дату j -го года; G_m – порядковый номер сезона. Процент данной

выборки от общего числа исходных данных (датасета) подбирается экспериментальным путем, как и в валидационной и проверочной выборках.

b) Валидационная (validation, 20%) и проверочная (test, 20%).

8. **Количество эпох.** Установлено значение 250 (подбиралось экспериментально по достижению наиболее точного результата). Диапазон выбора варьировался от 50 до 1 000 эпох. С количеством эпох $> 1\ 000$ ИНС переобучалась, что давало ее плохую локальную сходимость.

9. **Скорость обучения α** (0,90), подобранная экспериментальным путем с дальнейшим обучением искусственной нейронной сети [168].

10. **Вид нейрона** (классический или модифицированный – пункт 2.2.1).

Предлагаемый метод обучения искусственной нейронной сети второго поколения, архитектура которой представлена на рисунке 2.7, для прогнозирования потребления электроэнергии основан на методе Backpropagation [176]. Исходя из математических основ данного метода, при применении вышеобозначенных параметров необходимо нормализовать входные значения количества потребляемой энергии EE_{ji}^k для предотвращения исчезающих и взрывающихся градиентов следующим соотношением:

$$Ee_{ji}^k = \frac{EE_{ji}^k - Ee_{\min}}{Ee_{\max} - Ee_{\min}}, \quad (2.22)$$

где Ee_{\max} и Ee_{\min} – минимальное и максимальное значения количества потребляемой электроэнергии по всем данным, подаваемым на входной слой нейронной сети.

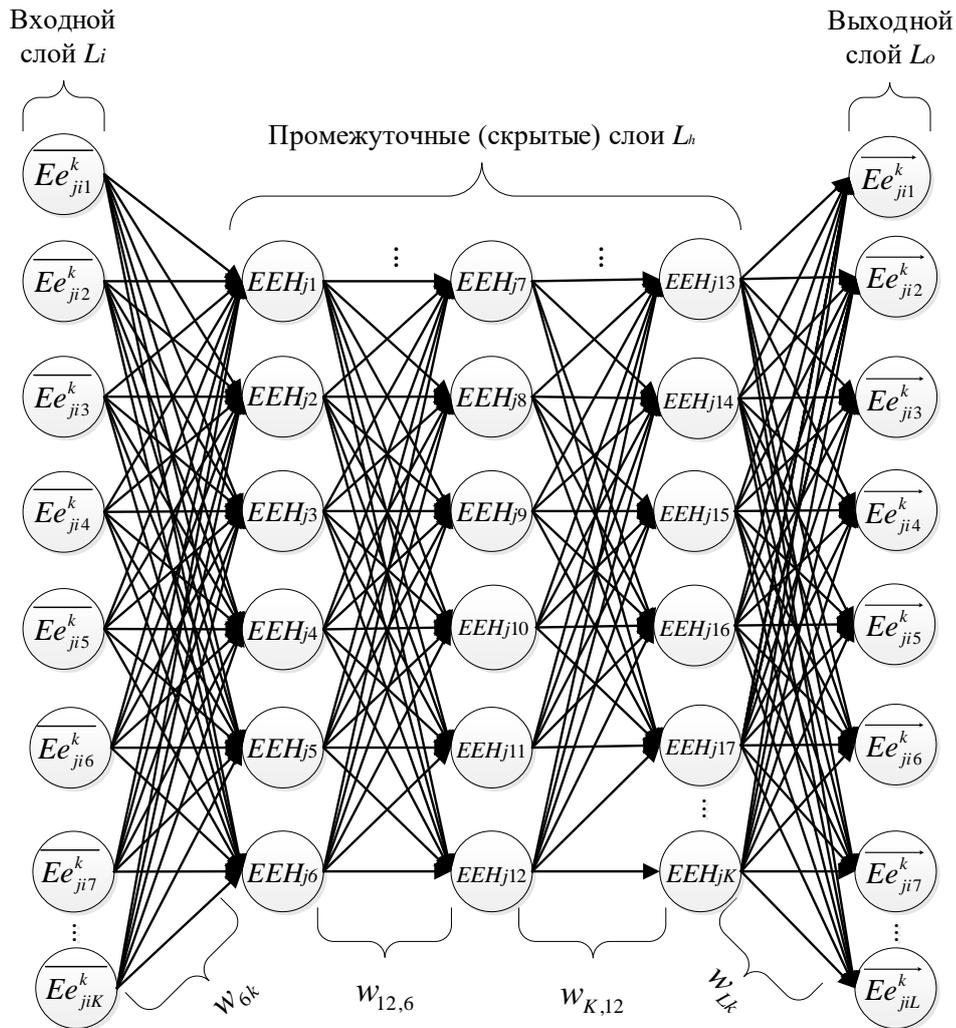


Рисунок 2.7 – Пример структуры нейронной сети второго поколения

На рисунке 2.7 для нумерации нейронов входного и промежуточного слоев введем индекс K , а для выходного – L . Подэтапы обучения нейронной сети на этапе предобработки данных (рисунок 2.6) представлены на рисунке 2.8.

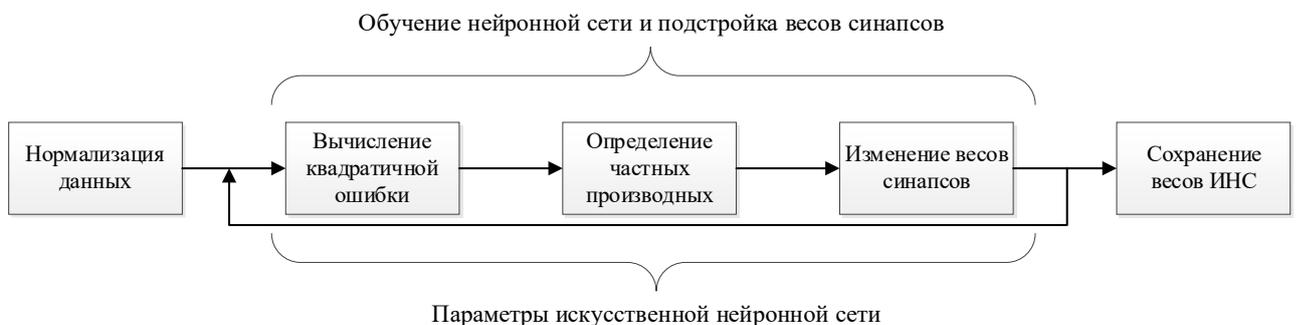


Рисунок 2.8 – Подэтапы обучения нейронной сети

Вычисление среднеквадратичной ошибки. Как правило, входы сети рассматриваются как входной вектор (\overline{EE}_{ji}^k) , где $\overline{EE}_{ji}^k = [\overline{Ee}_{ji1}^k, \overline{Ee}_{ji2}^k, \dots, \overline{Ee}_{jiK}^k]$, а выходы сети можно представить в качестве вектора выхода (\overline{EE}_{ji}^k) , где $\overline{EE}_{ji}^k = [\overline{Ee}_{ji1}^k, \overline{Ee}_{ji2}^k, \dots, \overline{Ee}_{jiL}^k]$, при этом $K = 1, \dots, M$, а $L = 1, \dots, N$, в которых M и N представляют собой размерность входного и выходного векторов. Соответственно, обучающая выборка представляет собой множество пар R выходных векторов \overline{Ee}_{ji}^k и желаемых (эталонных) выходных векторов \overline{Ee}_{ji}^k :

$$R = \{(\overline{Ee}_{ji1}^k, \overline{Ee}_{ji1}^k), (\overline{Ee}_{ji2}^k, \overline{Ee}_{ji2}^k), \dots, (\overline{Ee}_{jiL}^k, \overline{Ee}_{jiL}^k)\}. \quad (2.23)$$

Отсюда может быть произведен расчет фактического выхода \overline{Ee}_{jiL}^k выходного слоя:

$$\overline{Ee}_{jiL}^k = f(\sigma_L), \quad (2.24)$$

где σ_L – взвешенная сумма входов нейрона L , f – функция активации нейронов выходного слоя. Функция активации σ для нейронов промежуточных слоев, определяется по соотношению

$$\sigma(\sigma_{L_h}) = \frac{1}{1 + \exp(-\sigma_{L_h})} \quad (2.25)$$

и является функцией от взвешенной суммы входов нейрона EEH_{jl} в каждом промежуточном слое L_h . Тогда исходя из (2.19) и (2.20) можем вычислить номер слоя z для нейрона с номером l по формуле с целочисленным делением $z = \lfloor l / N_{il} \rfloor + 1$. Взвешенная сумма для нейронов промежуточных слоев и выходного слоя может быть представлена следующей формулой:

$$\sigma_{L_h} = \begin{cases} \sum_{n=1}^{N_{il}} (w_{nl} \overline{Ee}_{ji,n}^k), \text{ при } z = 1 \\ \sum_{n=(z-2)*N_{il}+l}^{(z-1)*N_{il}} (w_{nl} EEH_{jn}), \text{ при } z > 1 \end{cases}, \quad (2.26)$$

где w_{nl} – вес синапса (вес связи, соединяющей нейрон n нейроном l), l – нейрон в промежуточном слое.

Исходя из этого, мы можем определить среднеквадратичную ошибку для каждой пары векторов множества R путем суммирования среднеквадратичных ошибок в каждом выходном нейроне (2.27):

$$E_k^{ji} = \sqrt{\frac{1}{L} \sum_{l=1}^L (\overline{Ee_{ji,l}^k} - \overline{Ee_{ji,l}^k})^2}; \quad (2.27)$$

и как следствие, полную среднеквадратичную ошибку E путем суммирования всех E_k^{ji} для каждого значения всех k счетчиков ($k = 1, \dots, n$) и всех дат измерений в обучающей выборке (2.28):

$$E = \frac{\left(\sum_{k=1}^n \sum_j \sum_i E_k^{ji} \right)}{\left(n \cdot (\text{количество лет}) \cdot (\text{количество дат}) \right)}. \quad (2.28)$$

Таким образом, цель обучения состоит в минимизации E за счет нахождения соответствующего набора весов связи, соединяющего нейрон с нейроном (2.26).

Определение частных производных по весам синапсов. В качестве метода минимизации ошибки применим метод градиентного спуска, суть которого сводится к поиску минимума и максимума функции за счет движения вдоль вектора градиента, т.е. поиск минимума и максимума происходит за счет движения в направлении антиградиента. При этом градиент функции потерь ∇E представляет из себя вектор частных производных: $\nabla E = \left[\frac{dE}{dw_{11}}, \dots, \frac{dE}{dw_{nl}} \right]$, где $\frac{dE}{dw_{nl}}$ – частная производная функции ошибки по весу синапса. Количество частных производных соответствует количеству весов.

Исходя из вышеизложенного и соотношения (2.28) можем отметить, что

$$\frac{dE}{dw_{nl}} = \overline{Ee_{ji,l}^k} - \overline{Ee_{ji,l}^k}, \quad (2.29)$$

при этом, опираясь на (2.26), мы получаем

$$\frac{d\sigma_{L_h}}{dw_{nl}} = \sum_{n=(z-2)*N_{il}+1}^{(z-1)*N_{il}} EEH_{jn}. \quad (2.30)$$

Исходя из (2.29) и (2.30), определяем частную производную E по весу w_{nl} для выполнения в (2.31) градиентного спуска для весов синапсов между промежуточным и выходным слоями:

$$\frac{dE}{dw_{nk}} = \sum_{l=1}^L (\overline{Ee_{ji,l}^k} - \overline{Ee_{ji,l}^k}) \cdot \overline{Ee_{ji,l}^k} \cdot ((\overline{\sigma(Ee_{ji,l}^k)} - \overline{Ee_{ji,l}^k}) \cdot EEH_{jn}). \quad (2.31)$$

При этом, если рассматривать частный случай производной E между входным и промежуточными слоями, все выходы зависят от w_{nk} и частная производная будет находиться следующим образом:

$$\frac{dE}{dw_{nk}} = \sum_{l=1}^L (\overline{Ee_{ji,l}^k} - \overline{Ee_{ji,l}^k}) \cdot \overline{Ee_{ji,l}^k} \cdot ((\overline{\sigma(Ee_{ji,l}^k)} - \overline{Ee_{ji,l}^k}) \cdot (w_{nk} \cdot EEH_{jn})) \cdot ((1 - EEH_{jn}) \cdot \overline{Ee_{ji,l}^k}). \quad (2.32)$$

Таким образом, уравнения (2.29) и (2.30) дают все необходимые величины $\frac{dE}{dw_{nl}}$ для применения (2.31) и (2.32) в рамках градиентного спуска для всех весов нейронной сети.

Изменение весов синапсов. Каждый вес будет изменен на dw для уменьшения E :

$$w_{nl}(t+1) = w_{nl}(t) + \Delta w_{nl}(t), \quad (2.33)$$

где $w_{nl}(t)$ и $w_{nl}(t+1)$ – весовой коэффициент синапса, соединяющего эти нейроны,

на итерациях t и $t+1$ соответственно, $\Delta w_{nl}(t) = -\left. \frac{dE}{dw_{nl}} \right|_t$. В настоящем

диссертационном исследовании для увеличения скорости обучения нейронной сети, добавляется ранее обозначенный параметр α , а для повышения точности расчетов прогнозных значений задается специализированный коэффициент чувствительности φ , который был также обозначен ранее:

$$\Delta w_{nl}(t) = -\left. \frac{dE}{dw_{nl}} \right|_t + (\varphi \cdot (\alpha \cdot \Delta w_{nl}(t-1))). \quad (2.34)$$

Основной смысл изменения заключается в создании диапазона весов синаптических связей (весовых коэффициентов) $\varphi \in [0,1;0,9]$. Это позволит избежать резких колебаний значений в матрице весовых коэффициентов, что является необходимым условием повышения точности расчета прогнозных значений.

На третьем этапе производится непосредственно формирование прогнозной модели и расчет прогнозных значений, а на четвертом этапе – денормализация полученных прогнозных значений EE_{ji+l}^k (2.35) в связи с изначальной нормализацией данных ($EE_{ji}^k \in [0;1]$):

$$EE_{ji}^k = Ee_{ji}^k \cdot (Ee_{\max} - Ee_{\min}) + Ee_{\min} \cdot \quad (2.35)$$

2.3.1. Модификация модели классического нейрона под задачи прогнозирования временных рядов

За основу для модификации был выбран классический вариант нейрона (рисунок 2.9, на примере входного слоя), состоящий из входных значений, синаптических связей (синапсов), ячейки нейрона, аксона и выходного значения.

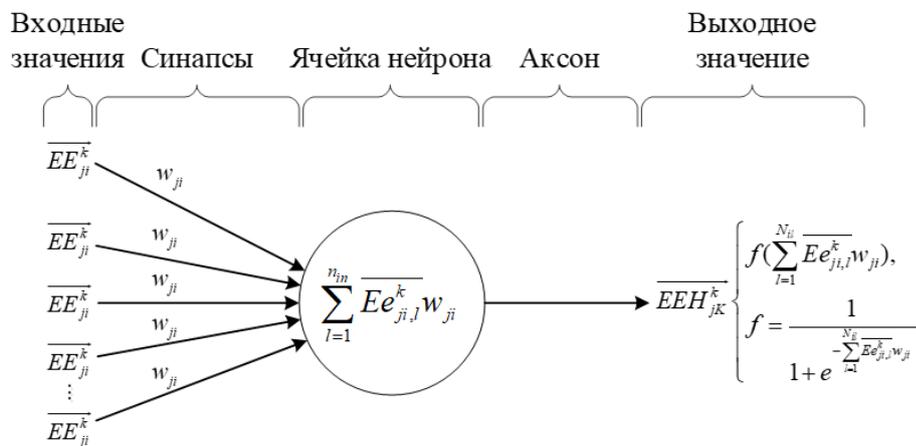


Рисунок 2.9 – Классический нейрон на примере входного слоя разработанной ИНС

На рисунке 2.9 n_{in} – число входов нейрона, $\overline{Ee_{ji}^k}$ – нормализованное значение количества потребляемой энергии, $\overline{EEH_{jK}^k}$ – выходное значение нейрона со входного слоя на промежуточный слой.

Модифицированный нейрон и его математическая модель показаны на рисунке 2.10.

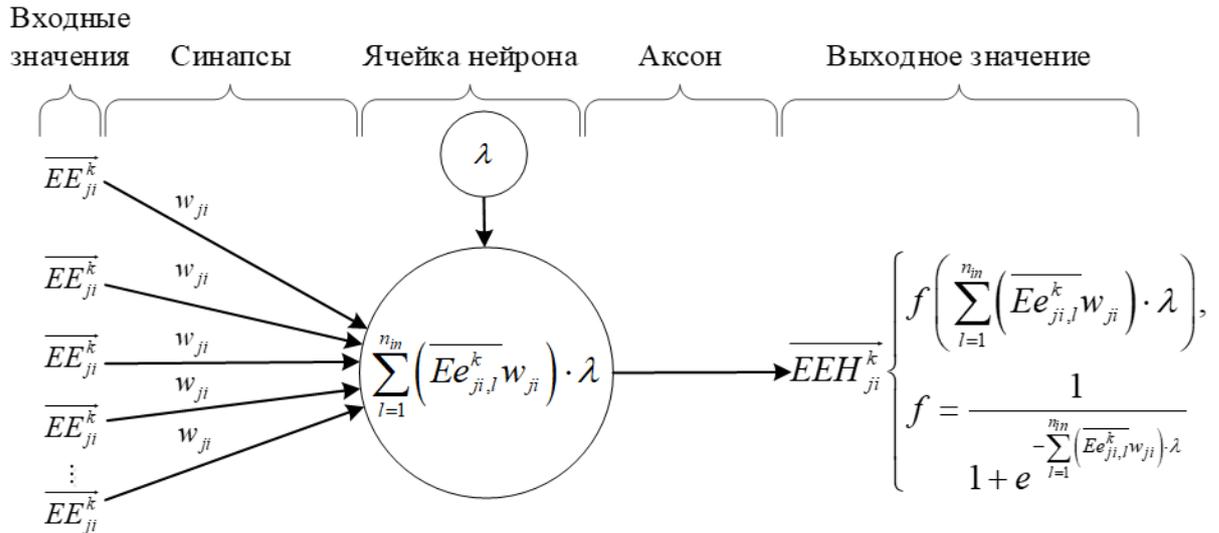


Рисунок 2.10 – Модифицированный нейрон с учетом добавления коэффициента реагирования на аномалии во временных рядах

Из рисунка 2.10 видно, что модифицированный нейрон имеет идентичную структуру (рисунок 2.9) за исключением добавления коэффициента реагирования $\lambda \in [\alpha_1, \alpha_2]$ (таблица 2.4) на аномалии во временных рядах. В качестве таких аномалий рассматриваются резкие всплески и понижения значений потребления электроэнергии, что может привести к неточным расчетам прогнозных данных. Сам коэффициент реагирования для каждого нейрона n рассчитывается по формуле (2.36) и был подобран в ходе проведения экспериментальной части (глава 4):

$$x_n = f \left(\sum_{l=1}^{n_{in}} (\overline{Ee_{ji,l}^k} w_{ji}) \cdot \lambda \right). \quad (2.36)$$

Для расчета коэффициента реагирования используем экспериментально $\lambda \in [\alpha_1, \alpha_2]$ определенные границы значений $\alpha_1 = 0,1$ и $\alpha_2 = 0,9$. При этом приняв в качестве стартового значения $\lambda = 0,1$ получаем формулу (2.36). Таким образом, рассчитываем среднее значение x_{sr} : $x_{sr} = \frac{1}{n} \sum_{i=1}^n x_i$, где n – количество нейронов. Следовательно, приняв это значение за математическое ожидание, задав некоторый интервал (обозначим как Δx , например, $\Delta x = 0,1x_{sr}$) вокруг данной величины, можем вычислить рассеивание, т.е. количество нейронов выпадающих за этот интервал $|x_i - x_{sr}| > \Delta x$. Если удельный вес таких нейронов больше наперед заданной величины (например, $\beta = 0,25$), необходимо увеличивать λ и повторять вычисления.

Таблица 2.4 – Значения коэффициента реагирования λ на аномалии во временных рядах

Значение коэффициента λ	Всплеск значений потребления электроэнергии во временном ряде, кВт*ч
0,1	50–119
0,2	120–179
0,3	180–239
0,4	240–289
0,5	290–399
0,6	400–899
0,7	900–4999
0,8	5000–8999
0,9	9000–15000

При помощи данного коэффициента мы учитываем такие всплески и понижения, что позволяет минимизировать последствия в виде неточности прогнозных значений. Таким образом, локальная сходимость модели ИНС повышается, как следствие, точность прогноза увеличивается, что подтверждается экспериментом, представленным в главе 4.

Выводы по главе 2

В ходе исследований и разработок были получены следующие результаты:

1. Разработаны метод и модель обработки информации, представленной в виде сетевого потока данных с применением импульсной нейронной сети. Данные метод и модель позволяют автоматически формировать исходный датасет для прогнозирования рекуррентной нейронной сетью, убирая из него лишний шум при помощи правил фильтрации, получаемых в ходе обучения импульсной нейронной сети. Обучение импульсной нейронной сети происходит на основе модифицированного метода обратного распространения ошибки. Исходный датасет формируется из следующих данных: дата измерения; № счетчика; потребляемая электроэнергия; температура воздуха; количество осадков; выходной день или нет; атмосферное давление и сезон. При этом данные с датчиков и внешних источников передаются в виде гиперкуба в импульсную нейронную сеть.

2. Разработаны метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети. Данные метод и модель позволяют выполнять раннее прогнозирование на 30 суток и состоит из следующих этапов: принятие решения о прогнозировании (т.е. ЛПП принимает решение по количеству дней, на которые будет осуществлен прогноз); предобработка данных; формирование прогнозной модели; постобработка данных. Этап предобработки данных заключается в нормализации данных полученного датасета, сформированного на основе импульсной нейронной сети. При этом данные уже присутствуют в БД MySQL и являются структурированными за счет решения задачи №1.

Таким образом, за счет разработанных в главе 2 методов и моделей стало возможным раннее и точное (с минимальным процентом отклонений, глава 4) прогнозирование потребления электроэнергии, что позволяет повысить точность

экспертных заключений (глава 3) и уменьшает роль человеческого фактора при принятии решений в различных ситуациях (глава 3).

ГЛАВА 3. РАЗРАБОТКА ПРОТОТИПА СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ОБРАБОТКИ БОЛЬШИХ ТЕМПОРАЛЬНЫХ ДАННЫХ И ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

В данной главе представлен прототип СППР на основе интеграции методов и моделей, разработанных в главе 2. Показаны базовая концепция реализации прототипа СППР, схема работы и входящие в состав системы поддержки принятия решения модули.

Основные функции и задачи, выполняемые прототипом СППР:

- информационная поддержка принятия управленческих решений на основе автоматизированной обработки больших разнородных и темпоральных данных;
- прогнозирование временных рядов и генерация правил фильтрации сетевого потока данных на основе автоматизированной обработки больших данных.

Разрабатываемое программное обеспечение позволяет свести к минимуму риски финансовых убытков предприятия за счет принятия решений ЛПР.

3.1. Разработка системной модели и четырехкаскадной структурной схемы управления с включением прототипа СППР

Предприятие (рисунок 3.1) подразделяется на три составляющих, в которых присутствуют стратегический и операционный уровни планирования: электростанции, магистральные и распределительные сети. На электростанциях производится планирование потребления электроэнергии и выделяемых на это мощностей, а также производство и передача электроэнергии в магистральные сети, учет и ОДУ, обслуживание и ремонт оборудования, сбыт электроэнергии с использованием системы менеджмента качества. Магистральные сети представляют собой сети, связывающие отдельные регионы, их крупнейшие источники и центры энергопотребления. По распределительным электрическим

сетям производится передача электроэнергии потребителям. В данном случае прототип СППР выступает инструментом для информационной поддержки принятия управленческих решений на оперативном уровне, к которым относятся планирование и проведение мероприятий по энергосбережению и резервирование электроэнергетических мощностей.

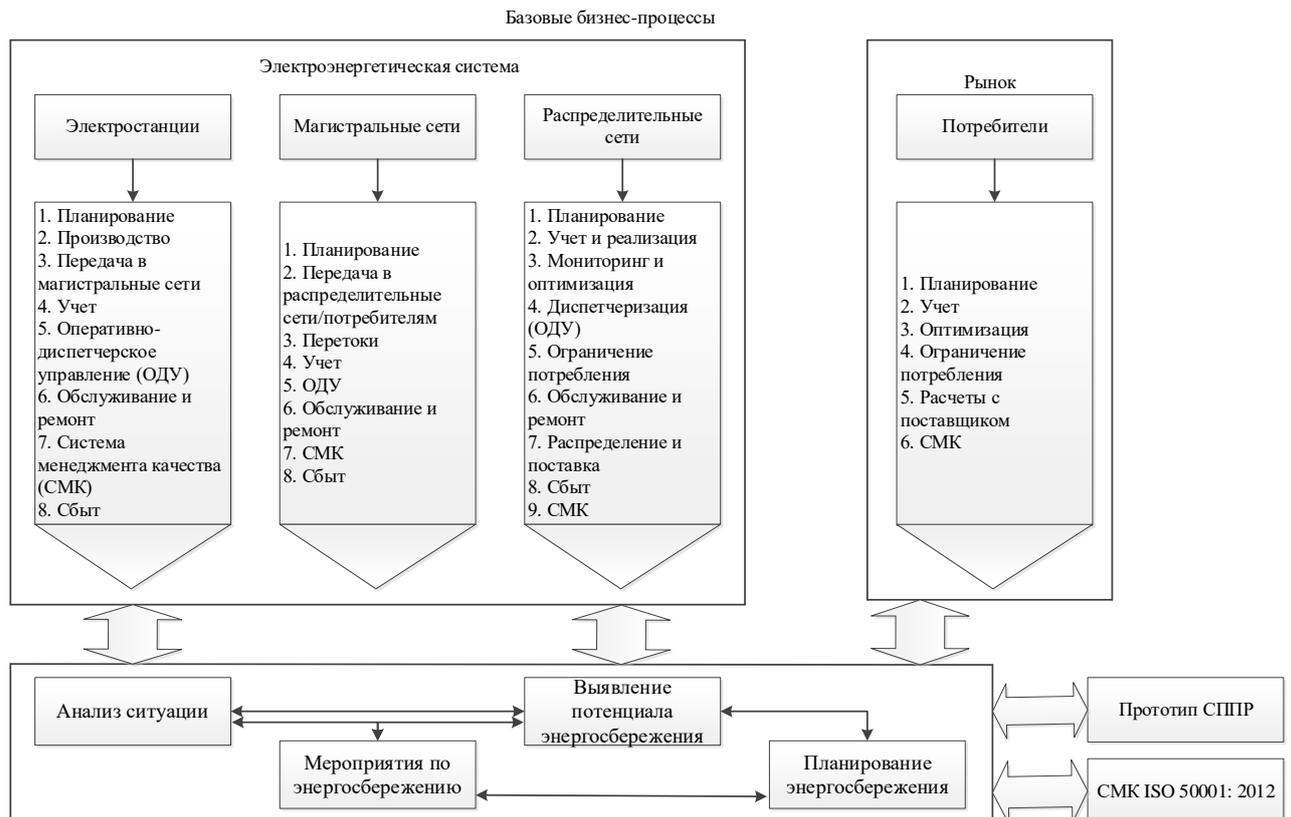


Рисунок 3.1 – Системная модель бизнес-процессов предприятия электроэнергетики на примере производственного отделения «Кумертауские электрические сети»

Более того, в связи с повсеместной компьютеризацией и автоматизацией, в том числе и на предприятиях электроэнергетики, появляются несколько информационных систем, конкурирующих между собой за сетевой поток данных, в рамках которого передаются данные для обучения ИмНС, входящей в состав прототипа СППР, со smart-датчиков («умных» датчиков) подстанций. Также на предприятии присутствуют датчики, взаимодействующие сразу с несколькими

информационными системами, которые в нужный момент времени получают эти данные в рамках корпоративной сети предприятия. Эти данные, курсирующие в рамках корпоративной сети предприятия, имеют большой процент зашумленности⁵, который необходимо отфильтровывать за счет правил фильтрации сетевого потока данных для формирования исходного датасета с целью обучения нейронных сетей. При этом сформированный единожды датасет не может считаться постоянно оптимальным по причине изменения внешней среды: отказ в работе датчиков; появляются новые данные, а также новые системы, перенаправляющие информационные потоки, что приводит к увеличению задержки при передаче потоков данных; изменение количества датчиков. Как следствие, необходимо анализировать сетевой поток данных на предмет необходимых данных для прогнозирования потребления электроэнергии. Таким образом, было предложено использовать два ранее разработанных метода, основанных на импульсной и рекуррентной нейронных сетях, а также их модификациях, что, в итоге, позволило повысить до 95% точность прогнозных значений потребления электроэнергии.

Предлагаемое решение ориентировано на информирование ЛПП о возникших ситуациях. При возникновении ситуации, предусматривающей возможность информационной поддержки принятия решения ЛПП, предлагаются варианты на выбор для исправления сложившейся проблемы. В системе предусмотрено вычисление прогнозных показателей на требуемый период, которые могут быть использованы для решения стратегических задач (например, принятие решения о необходимых объемах резервирования энерго мощностей). Данные возможности предлагаемого решения определяют получение (вычисление) ожидаемой эффективности.

Предлагаемые методы и модели апробировались путем публикаций в научно-технических журналах [164–169], практической реализации [177] и регистрации

⁵ Информация по зашумленности основывается на проведенных экспериментах по фильтрации данных

разработанных программных продуктов (приложение А). В рамках текущей диссертации предлагаемый прототип СППР направлен на информационное сопровождение принятия решений на основе обработки больших разнородных и темпоральных данных и прогнозирования временных рядов в электроэнергетике, при этом результаты апробации представлены в главе 4 и подробнее рассмотрены в рамках разработки контура управления процессом.

Контур управления представляет собой замкнутую, имеющую прямую и обратную связи, цепь (совокупность) звеньев (элементов) системы управления, посредством которой происходит соединение субъектов и объектов управления. В прототипе СППР объектом управления является информационная поддержка принятия управленческих решений на основе процесса обработки больших разнородных и темпоральных данных на примере внешнего сетевого потока данных и прогнозирования временных рядов потребления электроэнергии.

В качестве управления процессом информационной поддержки принятия управленческих решений представлены четыре каскада:

– контур обучения. В рамках контура обучения выполняется тренировка импульсной и рекуррентной нейронных сетей на основе сетевого потока данных и значений потребления электроэнергии;

– контур адаптации. В рамках контура адаптации происходит подстройка прототипа СППР к текущей ситуации;

– контур регулирования. В рамках контура регулирования выполняется переключение прототипа СППР между ситуациями Δ_1 , Δ_2 , Δ_3 , Δ_4 ;

– контур автоматического регулирования.

Контур управления процессом информационной поддержки принятия управленческих решений на основе обработки больших разнородных и темпоральных данных и прогнозирования временных рядов представлен на рисунке 3.2.

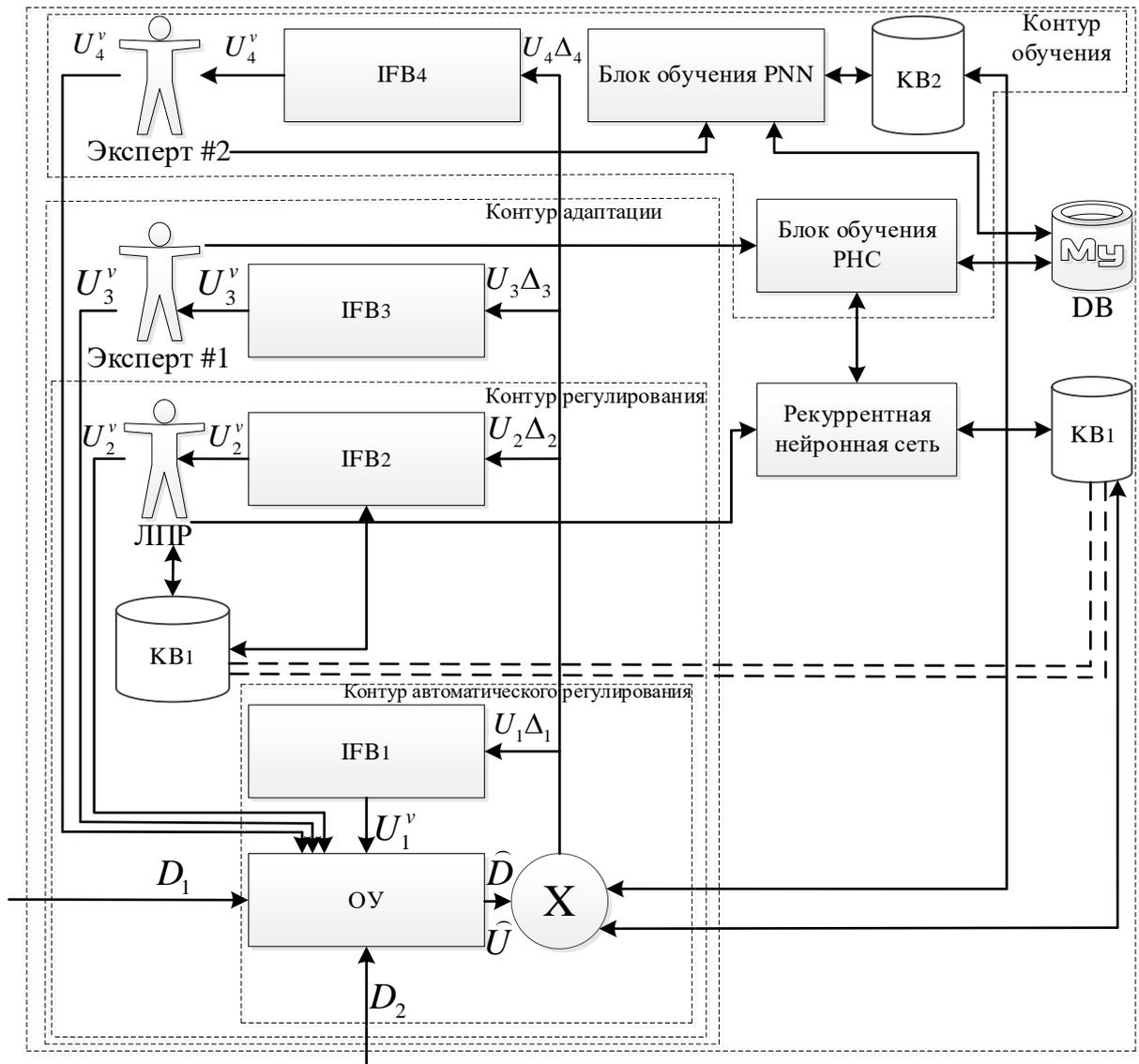


Рисунок 3.2 – Контур управления функциональным процессом прогнозирования потребления электроэнергии с включением системы информационной поддержки принятия решений

На рисунке 3.2 введены следующие обозначения:

1. ОУ – объект управления представленной системы информационной поддержки принятия управленческих решений. Управленческие решения могут быть связаны с:

– резервированием дополнительных мощностей электроэнергетического предприятия на основе прогнозных значений потребления электроэнергии, в том числе и с проведением энергосберегающих мероприятий;

– внедрением правил фильтрации сетевого потока данных для решения задач оптимизации сети различных электроэнергетических организаций, а также центров обработки данных.

2. Эксперт #1 – разработчик, эксперт #2 – системный администратор.

3. $IBF_1, IBF_2, IBF_3, IBF_4$ – блоки формирования воздействия.

4. DB (Database) – база данных MySQL.

5. KB_1, KB_2 – базы знаний, получаемые в результате обработки ретроспективных данных.

6. Δ_1 – ситуация, не требующая вмешательства ЛПП: случай, когда выходные данные удовлетворяют ЛПП/экспертов.

7. Δ_2 – ситуация, требующая вмешательства ЛПП, но не требующая подключения эксперта #1: случай, когда требуются необходимые правила фильтрации или прогнозные значения для принятия управленческих решений.

8. Δ_3 – ситуация, требующая подключения эксперта #1: случай, когда требуются прогнозные значения потребления электроэнергии для принятия управленческих решений, при этом параметры ИНС настроены некорректно, как следствие, меньшая точность значений на выходе.

9. Δ_4 – ситуация, требующая подключение эксперта #2 в двух случаях: при существенном изменении сетевого потока данных (внешнее возмущающее воздействие); при некорректном формировании правила обработки информации (сетевого потока данных).

10. $\hat{D} = \{D_1, D_2\}$, где D_1 – внешний сетевой поток данных и ретроспективные данные, D_2 – внешнее возмущающее воздействие.

11. $\hat{U} = U_1^v, U_2^v, U_3^v, U_4^v$ – варианты управляющего воздействия на ОУ, где v – текущая ситуация, при этом $v = 1, \dots, n$, где n – количество ситуаций.

12. U_1 – информационный вектор управляющих воздействий после срабатывания переключателя ситуаций Δ .

13. U_2 – информационный вектор управляющих воздействий после переобучения рекуррентной нейронной сети.

14. U_3 – информационный вектор управляющих воздействий после переобучения импульсной нейронной сети.

15. U_4 – информационный вектор управляющих воздействий при необходимости принятия управленческих экспертом #2.

Для решения задач обработки больших разнородных и темпоральных данных на примере сетевого потока данных и прогнозирования временных рядов потребления электроэнергии прототип СППР должен обеспечивать следующими функциями:

1. Автоматизированный сбор и нормализация ретроспективных данных для обучения PNN и РНС. Данная функция реализуется в подсистеме предобработки и постобработки данных (рисунки 2.2, 2.4) за счет методов нормализации (2.7, 2.22) и денормализации (2.8, 2.35) информации, что сведено в единую структурную схему системы предобработки и постобработки данных в рамках прототипа СППР на рисунке 3.3.

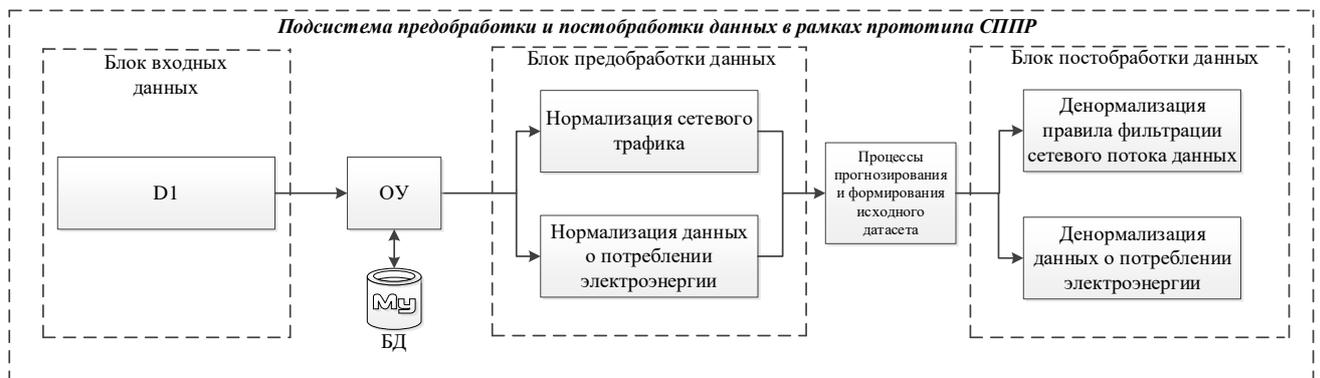


Рисунок 3.3 – Структурная схема системы, входящей в РНС и PNN, предобработки и постобработки данных в рамках прототипа СППР

2. Формирование правил фильтрации сетевого потока данных на выходе импульсной нейронной сети с их последующим внесением в базу данных MySQL.

3. Прогнозирование временных рядов на примере значений потребления электроэнергии.

4. Переключение между ситуациями, когда необходимо задействовать экспертов или ЛПР в различных ситуациях, что было показано на рисунке 3.2.

Учитывая, что прототип СППР ориентирован на информационное сопровождение поддержки принятия решения, отдельные модели процессов и функций организации не описывались. Информационная модель, связанная с обработкой данных для информационной поддержки принятия решений, представлена на рисунке 3.4 и полностью описана в разделе 3.2. Функциональная структура предложенной системы в привязке к рисунку 3.1 представлена в разделе 3.3. Основные возможности и ограничения предложенного решения представлены в пункте 3.3.1.

Обеспечение полноты своевременности и достоверности сведений ограничивается возможностями сбора сведений данных для использования нейронной сети. Дискретизация полученных данных определяет возможность своевременности принятия решений. Учитывая ограниченную, но в то же время высокую 92–95% (глава 4) точность прогнозных значений в рамках предлагаемого прототипа СППР, и отнесение объектов электроэнергетики к критической инфраструктуре, речь идет только об упреждающем информировании ЛПР о возникших ситуациях.

3.2. Разработка информационной модели прототипа СППР

Информационная модель обработки больших разнородных и темпоральных данных и прогнозирования временных рядов построена в соответствии с методологией IDEF1X (рисунок 3.4).

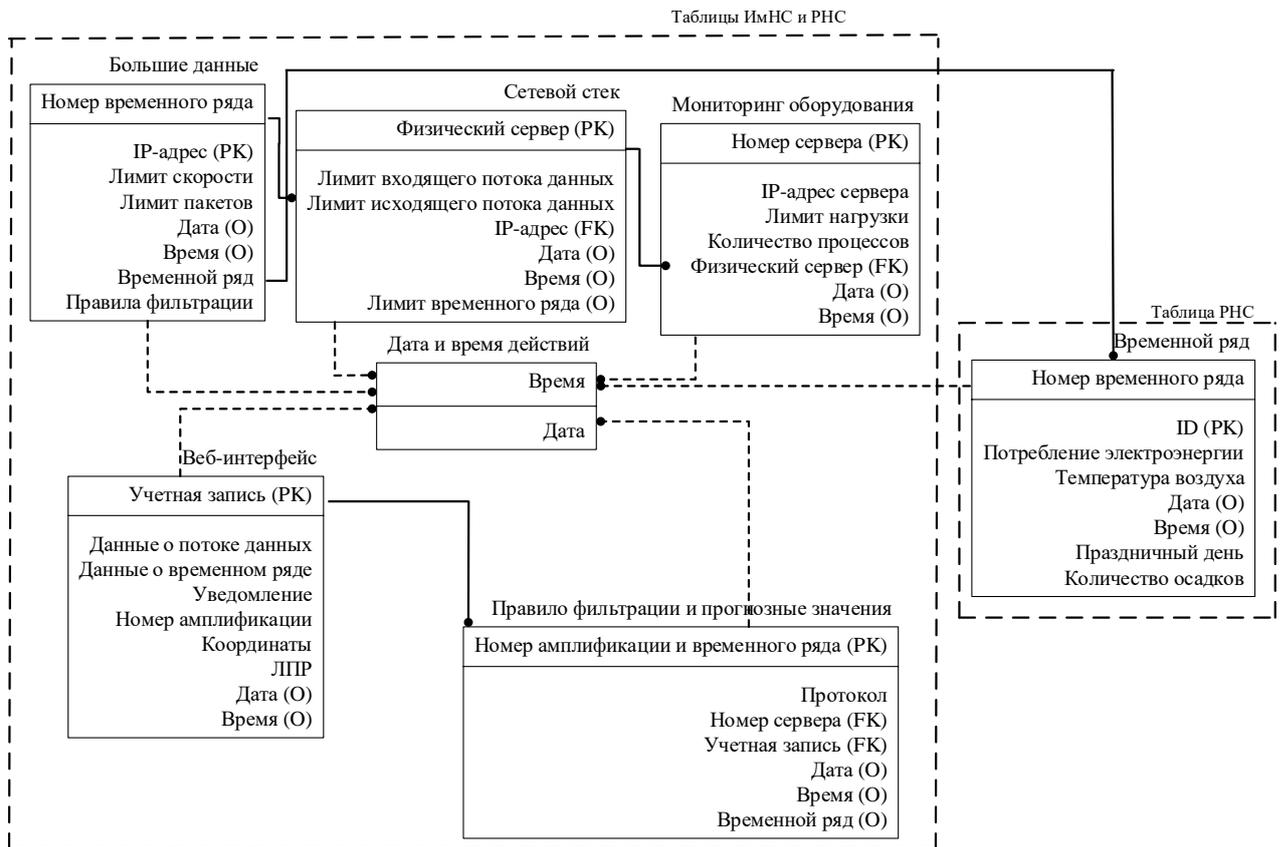


Рисунок 3.4 – Информационная модель (фрагмент) базы данных MySQL прототипа СППР

Фрагмент информационной модели (физическое представление базы данных, см. рисунок 3.3) представляет собой связь созданных сущностей методом «один – ко – многим». Сущности разработанной базы данных:

- большие данные;
- сетевой стек;
- мониторинг оборудования;
- правило фильтрации и прогнозные значения;
- веб-интерфейс;
- дата и время действий;
- временной ряд.

Связи между сущностями позволяют описать механизм наследования атрибутов таблиц в базе данных. Таким образом, в таблицу «Большие данные» происходит запись следующей информации: информация о сетевом потоке данных

(IP-адрес, лимит скорости, лимит пакетов, дата, время); сетевой поток данных в виде временных рядов в области электроэнергетики для последующей работы искусственных нейронных сетей, а также вывода данных в веб-интерфейс; правила фильтрации. В таблицу «Сетевой стек» происходит запись лимитов (пропускная способность сетевого стека) входящего/исходящего сетевого потока данных, в том числе и временного ряда, а также записываются отредактированные параметры сетевого стека и номер физического сервера в атрибут «Физический сервер». В таблицу «Мониторинг оборудования» происходит запись номера физического сервера в атрибут «Номер сервера», также записываются: «IP-адрес сервера» (для идентификации), «Лимит нагрузки» (для определения порога нагрузки на каждом сервере перед распределением), «Количество процессов» (для определения степени загруженности ЭВМ) и «Физический сервер» (для определения конфигурации сервера). В таблицу «Правила фильтрации и прогнозные значения» заносятся аномалии сетевого потока данных (изменение вида и типа сетевого потока данных) по номерам (атрибут «Номер амплификации»), а также записываются: «Временной ряд», «Протокол» (для определения протокола сетевого потока данных через веб-интерфейс), «Номер сервера» (для определения сервера в кластере, на который подается входящий сетевой поток данных), «Учетная запись» (для определения IP-адреса объекта в веб-интерфейсе). В таблице «Веб-интерфейс» заполняются атрибуты: «Учетная запись» (для авторизации в административной панели управления АПК), «Данные о сетевом потоке данных» (для генерации мониторинга скорости внешнего сетевого интерфейса), «Данные о временном ряде» (для вывода данных о временном ряде в веб-интерфейс), Уведомление (для сохранения текста отправленного уведомления), «Номер амплификации» (для вывода статистики амплификаций на все физические серверы), а также «координаты» (для определения стран, из которых поступает сетевой поток данных) и «ЛПР» (учетная запись). В таблицу «Временные ряды» записываются: ID, значения потребления электроэнергии, дата и время, праздничные и выходные дни, а также количество осадков.

3.3. Требования по разработке прототипа СППР, архитектура, разработка и основной функционал

Требования, предъявляемые к прототипу СППР, соответствуют ГОСТР 52611–2006 «Средства информационной поддержки жизненного цикла продукции. Безопасность информации» и ГОСТР ИСО/МЭК 15408-1-2012 «Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий» и являются следующими:

1. Функциональные характеристики. Система поддержки принятия решений должна представлять совокупность методических и программных средств решения задач:

1.1. Автоматическое формирование исходного датасета на основе интеллектуальных технологий, в том числе и получение метеорологических данных с внешних источников.

1.2. Прогнозирование потребления электроэнергии несколькими методами, т.е. должен быть выбор метода прогнозирования.

1.3. Графическое отображение счетчиков, по которым будет выполняться прогноз значения потребления электроэнергии.

1.4. Сохранение результата в файл, в том числе и прогнозной ошибки для проведения сравнительного анализа с другими методами.

1.5. Вход в систему должен быть доступен после регистрации пользователя.

2. Используемые методы прогнозирования. Для прогнозирования потребления электроэнергии должны использоваться методы машинного и глубокого обучения.

3. Требования к методическому обеспечению. Методическое обеспечение должно быть реализовано в пользовательском интерфейсе системы, который должен предполагать: регистрацию и авторизацию пользователя; выбор метода прогнозирования; парсинг данных; формирование и сохранение исходного датасета в базу данных, а также прогнозных значений потребления электроэнергии в БД MySQL.

4. Требования к надежности согласно ГОСТ 27.002–2021, т.е. система должна сохранять во времени в установленных пределах значения всех параметров, необходимых для прогнозирования, в том числе и характеризующих способность системы выполнять требуемые функции в режиме авторизации. В этой связи необходимо:

4.1. Предусмотреть контроль вводимой информации и блокировку некорректных действий пользователя при работе с системой.

4.2. Обеспечить корректное завершение вычислений прогнозных значений при превышении имеющихся вычислительных ресурсов.

5. Требования к составу и параметрам технических средств. Система должна быть облачной, т.е. все вычисления производятся на выделенном физическом сервере/кластере, а управление должно осуществляться через браузер персонального компьютера и мобильного устройства любой операционной системы.

6. Требования к совместимости. Прототип СППР должен быть кроссплатформенным.

Прототип СППР необходим для оперативного и эффективного информационного сопровождения при принятии управленческих решений, в нашем случае – на примере обработки и фильтрации больших объемов сетевого потока данных, а также прогнозирования временных рядов для электроэнергетической области. Для этого модули прототипа СППР реализованы с учетом возможности работы всей системы поддержки принятия решений в двух режимах: с частичным участием человека (принятие наиболее важных решений) и без человека. Как результат – все модули прототипа СППР будут взаимодействовать между собой автоматически (без участия человека), а ЛПР необходимо вмешиваться только в критических ситуациях. Полная автоматизация необходима для обеспечения качественной обработки больших данных и прогнозирования временных рядов в кратчайшие сроки.

Необходимо отметить, что функциональная структура (архитектура) прототипа СППР подразделяется на несколько составляющих:

1. Обеспечение автоматического взаимодействия между каждым модулем и веб-частью.

2. Реализация платформы для бесконфликтной и взаимосвязанной работы модулей между собой.

3. Создание связей между базой данных MySQL и модулями прототипа СППР для дальнейшего хранения передаваемой информации.

4. Обработка и фильтрация больших объемов сетевого потока данных, а также прогнозирование временных рядов для поддержки принятия решений в области электроэнергетики.

Обобщенная функциональная структура прототипа СППР (архитектура) представлена на рисунке 3.5.

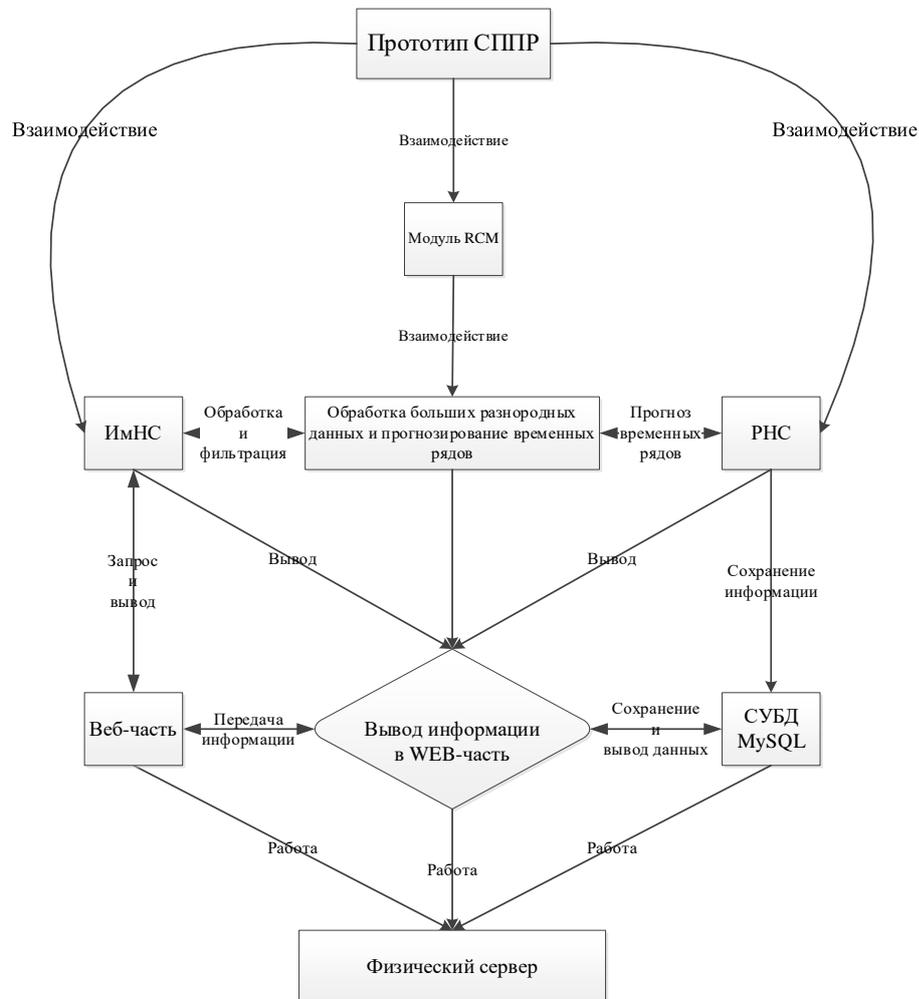


Рисунок 3.5 – Обобщенная функциональная структура (архитектура) прототипа СППР

Далее представлены алгоритмы работы предлагаемого прототипа СППР с использованием разработанных методов и моделей:

1. Взаимодействие модуля удаленного клиента (RCM) с искусственными нейронными сетями и веб-частью сведены в единую структурную схему (рисунок 3.6).

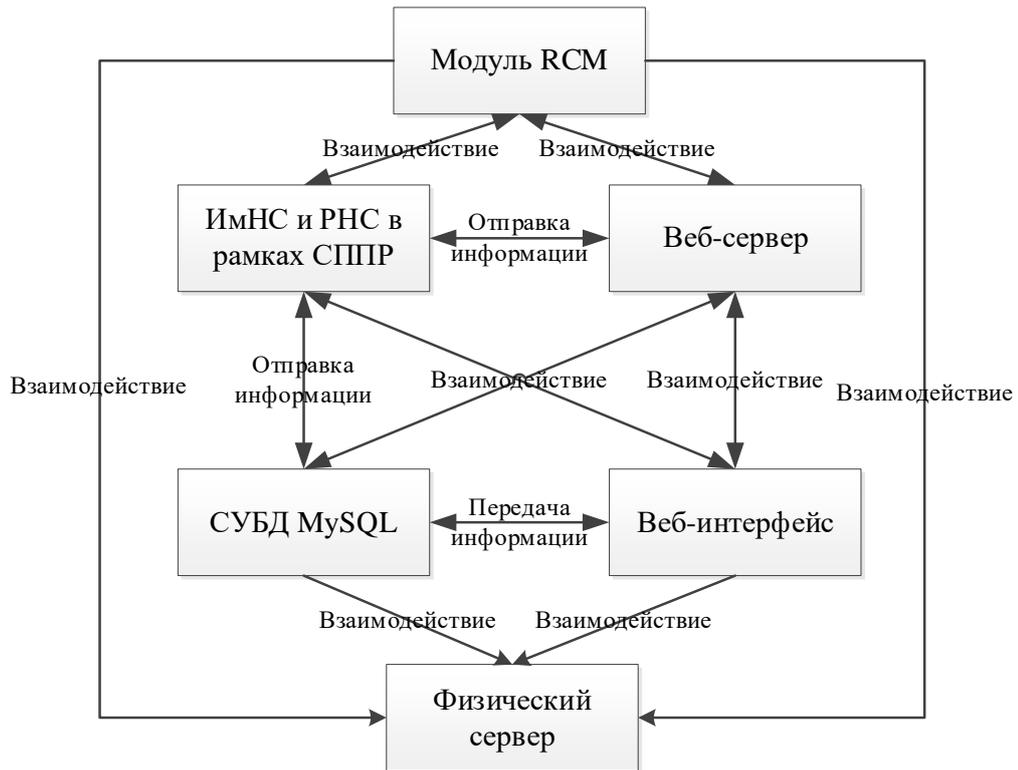


Рисунок 3.6 – Структурная схема взаимодействия между RCM, ИМНС, РНС и веб-интерфейсом

Удаленный клиент RCM производит управление всеми искусственными нейронными сетями (разработанные ранее методы и модели) посредством языков программирования Bash (усовершенствованная и модернизированная вариация командной оболочки Bourne shell) и Python. Благодаря данному модулю становится возможен вывод данных в веб-интерфейс.

Структурная схема взаимодействия ИМНС и РНС в рамках прототипа СППР представлена на рисунке 3.7.

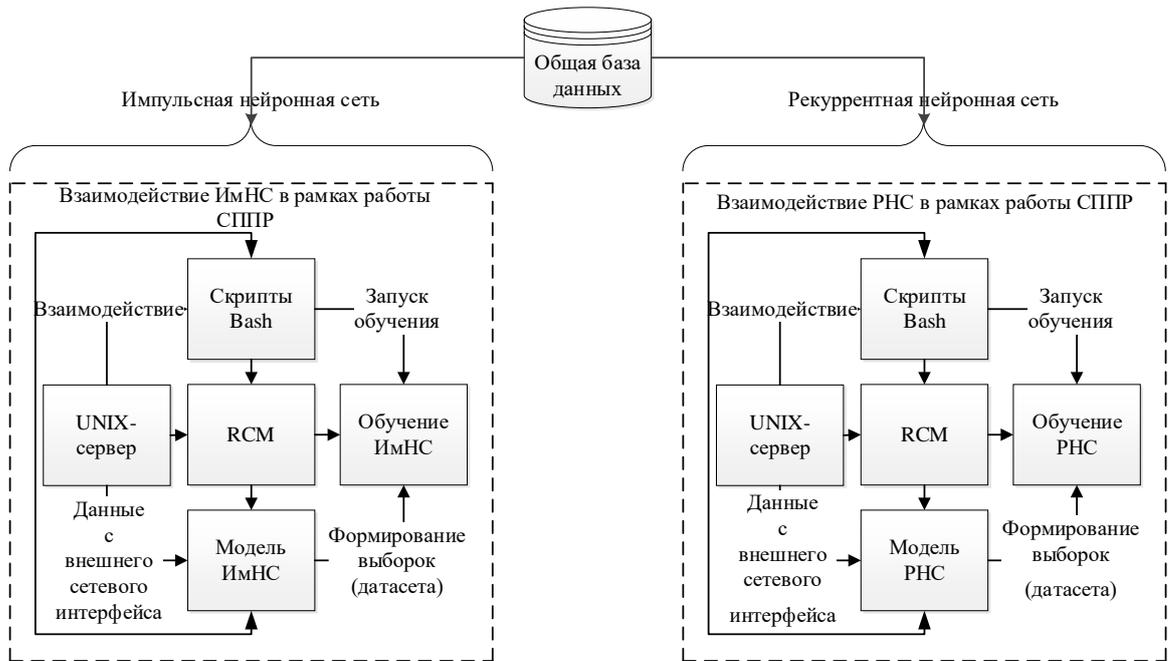


Рисунок 3.7 – Структурная схема работы и взаимодействия ИмНС, РНС и RCM

При входе в прототип СППР мы попадаем на главное окно входа, что показано на рисунке 3.8.

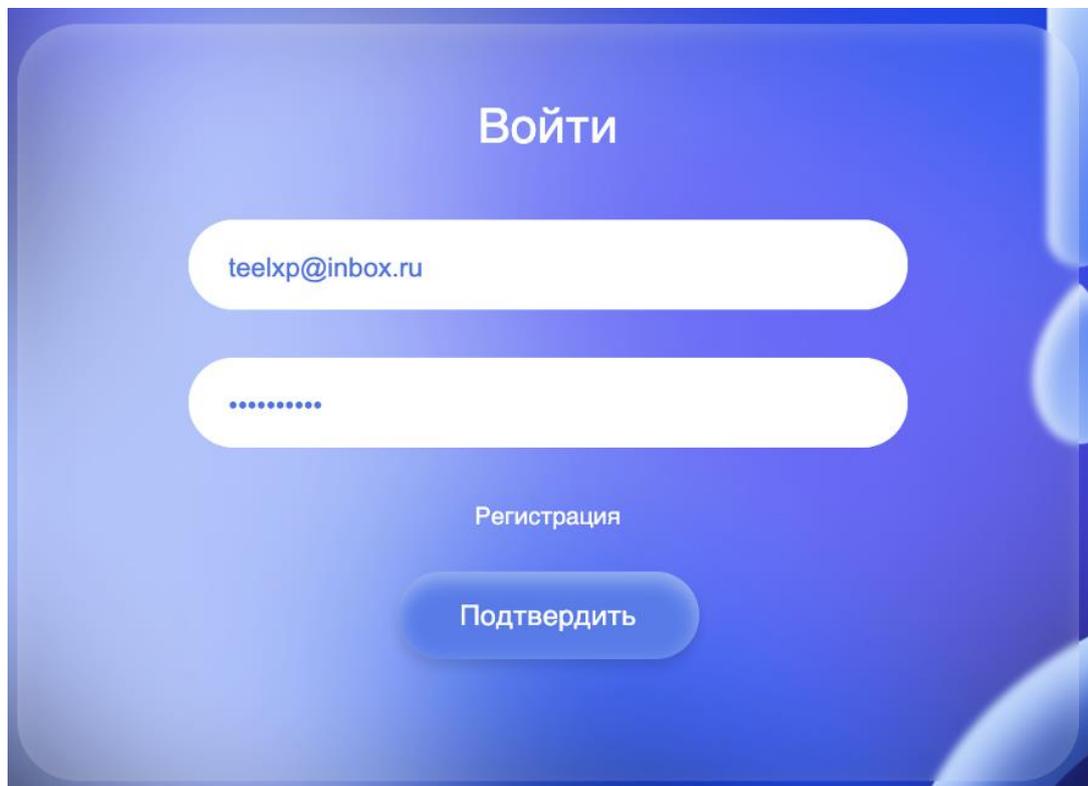


Рисунок 3.8 – Главное окно входа

Управление прототипом СППР, в том числе и разработанными ранее методами и моделями, реализованными в виде искусственных нейронных сетей (ИмНС и РНС), представлено на рисунках 3.9–3.10. Управление нейронными сетями происходит за счет модуля RCM, отвечающего за выполнение скриптов по запуску/остановке/обучению/перезагрузке (переобучению) и выполнению прогноза ИНС. Добавление в базу данных в обоих случаях происходит за счет модуля обработки больших разнородных и темпоральных данных, встроенного в искусственные нейронные сети (рисунок 3.11). В рамках РНС данный модуль подразделяется на подмодули: «MySQL» (рисунок 3.12), «ReadFile» (рисунок 3.13) и «Normalization» (рисунок 3.14). В ИмНС структура, архитектура и исходный код данного модуля являются идентичным, отличие заключается в том, что в модуль «Normalization» добавляется кодировщик данных, позволяющий создавать правила фильтрации сетевого потока данных.

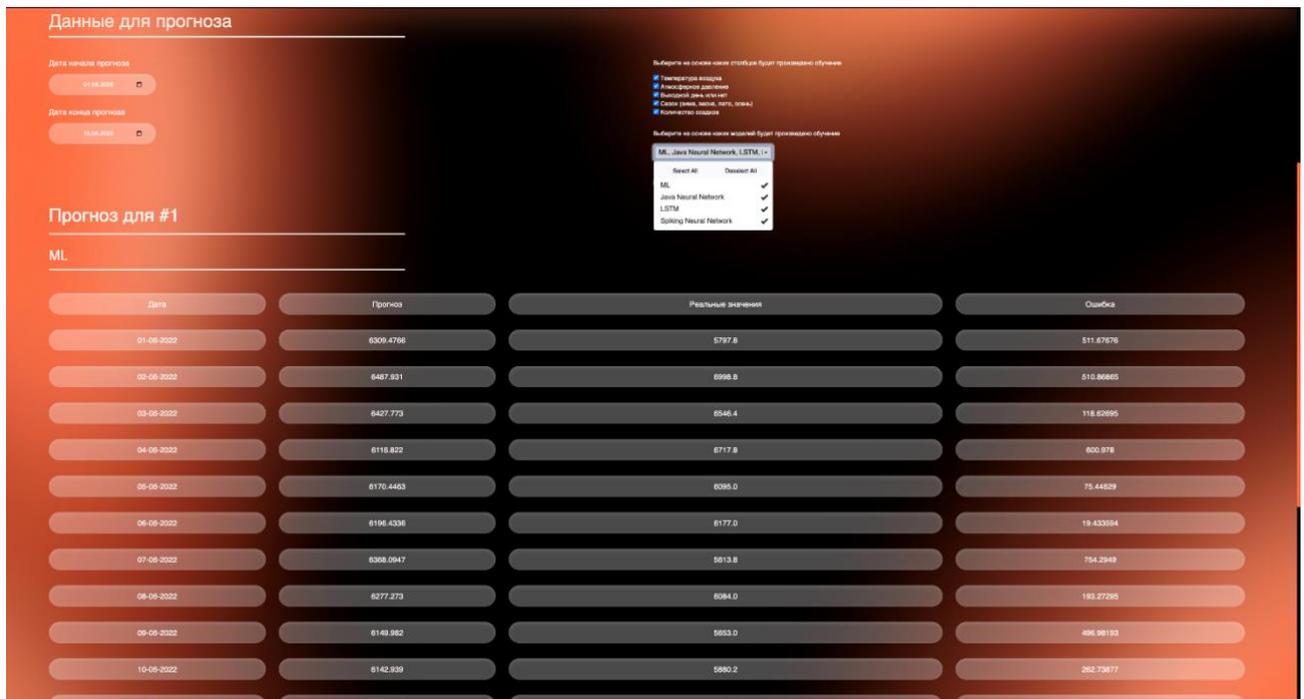


Рисунок 3.9 – Страница работы с РНС и ИмНС

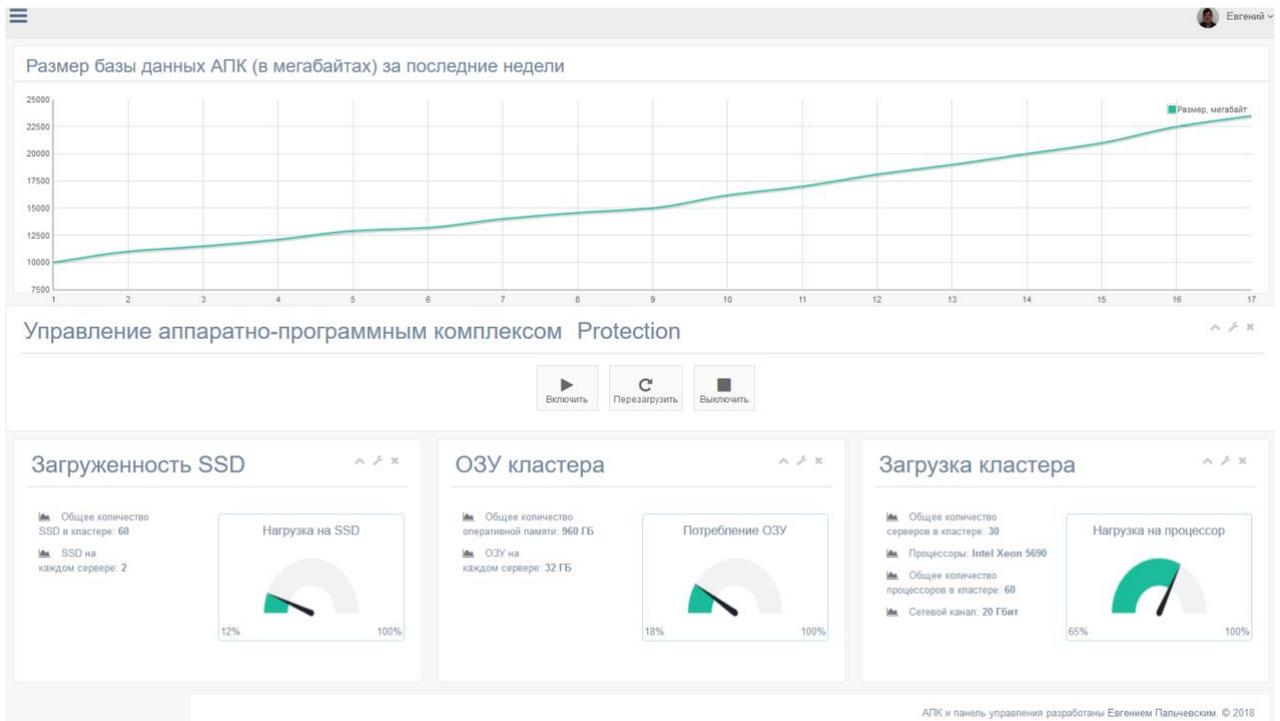


Рисунок 3.10 – Страница работы с ИмНС (мониторинг и управление прототипом СППР)

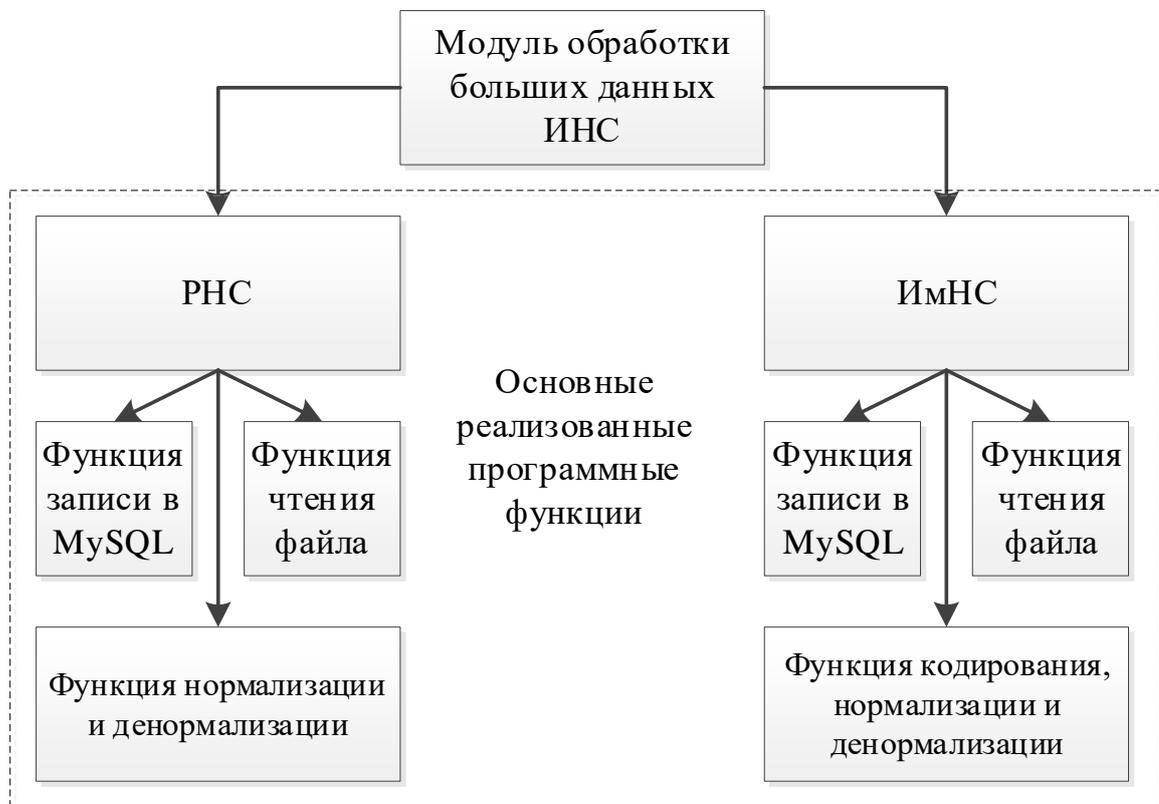


Рисунок 3.11 – Структура модуля обработки больших данных в РНС и ИмНС

```

1- import java.util.Date;
2
3 @Data
4 @Entity
5 @AllArgsConstructor
6 @Table(name = "complete_with_electro_sorted")
7- public class Energy {
8
9     @Id
10    private int id;
11
12    @Column(name = "Дата")
13    @DateTimeFormat(pattern = "yyyy-MM-dd")
14    private Date date;
15
16    @Column(name = "Потребляемая электроэнергия, кВт*ч")
17    private float electricityConsumption;
18
19    @Column(name = "Температура воздуха")
20    private int temperature;
21
22    @Column(name = "Количество осадков")
23    private float precipitationAmount;
24
25    @Column(name = "Выходной день или нет")
26    private int dayOffOrNot;
27
28    @Column(name = "Атмосферное давление")
29    private int atmospherePressure;
30
31    @Column(name = "Сезон (зима, весна, лето, осень)")
32    private int season;
33
34    @Transient
35    @DateTimeFormat(pattern = "yyyy-MM-dd")
36    private Date lastDate;
37
38-    public Energy() {
39    }
40
41 }
42

```

Рисунок 3.12 – Фрагмент исходного кода модуля обработки данных, отвечающего за взаимодействие модуля с MySQL

```

26 cursor = cnx.cursor()
27 cursor.execute(f'SELECT * FROM complete_with_electro_sorted_reserve;')
28
29 dataset = pd.DataFrame(cursor.fetchall(), columns=cursor.column_names)
30 dataset['Потребляемая электроэнергия, кВт*ч'] = dataset['Потребляемая электроэнергия, кВт*ч'].astype(np.float32)
31 dataset['Дата'] = dataset['Дата'].astype('string')
32
33- clear_dataset = dataset.drop(columns=["id", "Дата", "% счетчика", 'Прогнозные значения температуры воздуха',
34-                                     'Прогноз количества осадков']).astype(np.float32)
35 cursor.close()
36 cnx.close()
37
38 return dataset, clear_dataset
39
40
41 def make_inputs_and_targets(days_out, clear_dataset):
42 num_days_in, num_days_out = days_out * 2, days_out
43 inputs, targets = [], []
44
45 for i in range(len(clear_dataset) - num_days_in - num_days_out):
46     flag = True
47     for value in clear_dataset['Потребляемая электроэнергия, кВт*ч'].iloc[i:i + num_days_in + num_days_out]:
48         if value >= 14000 or value <= 3500:
49             flag = False
50     if flag:
51         inputs.append(list(itertools.chain(*clear_dataset.iloc[i:i + num_days_in].values.tolist()))))
52         targets.append(list(
53             clear_dataset['Потребляемая электроэнергия, кВт*ч'].iloc[i + num_days_in:i + num_days_in + num_days_out]
54         ))
55 return inputs, targets
56
57
58 def find_best_degree(X_train, X_test, y_train, y_test):
59 regression = LinearRegression()
60 best_degree, best_result = 1, 0
61
62 for degree in range(1, 5):
63     poly = PolynomialFeatures(degree=degree)
64     train_poly_features = poly.fit_transform(X_train)
65     test_poly_features = poly.fit_transform(X_test)
66
67     regression.fit(train_poly_features, y_train)
68     score = regression.score(test_poly_features, y_test)

```

Рисунок 3.13 – Фрагмент исходного кода модуля обработки данных, отвечающего за взаимодействие с файлом исходных выборок и восстановление данных

```

1 package org.neuralNetworks.demo.network;
2
3 import java.util.ArrayList;
4 import java.util.List;
5
6 public class Normalization {
7     public float[][] inputs;
8     public static List<double[]> minMax = new ArrayList<>();
9     public static List<Float> averages = new ArrayList<>();
10    public static List<Float> deviations = new ArrayList<>();
11
12    public void findMinMax(float[][] inputs) {
13        float[][] tmp = MatrixActions.transpose(inputs);
14        for (float[] floats : tmp) {
15            float max_value = MatrixActions.find_max_in_array(floats);
16            float min_value = MatrixActions.find_min_in_array(floats);
17            minMax.add(new double[]{max_value, min_value});
18        }
19    }
20
21    public float[][] normalize(float[][] inputs) {
22        float[][] tmp = MatrixActions.transpose(inputs);
23        for (int i = 0; i < tmp.length; i++) {
24            float max_value = (float) minMax.get(i)[0];
25            float min_value = (float) minMax.get(i)[1];
26            for (int j = 0; j < tmp[i].length; j++)
27                tmp[i][j] = (tmp[i][j] - min_value) / (max_value - min_value);
28        }
29        return MatrixActions.transpose(tmp);
30    }
}

```

Рисунок 3.14 – Фрагмент исходного кода модуля обработки данных, отвечающего за нормализацию/денормализацию данных

Таким образом, разрабатываемый прототип СППР может быть представлен структурными схемами (рисунки 3.3, 3.5–3.7), связанными единым контуром управления (рисунок 3.2).

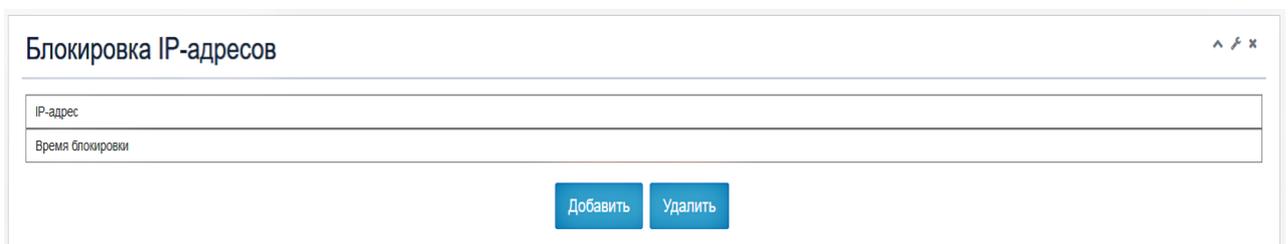
3.3.1. Основные возможности и ограничения ИмНС – метода и модели обработки сетевого потока данных в рамках прототипа СППР

1. Разработанный метод (ИмНС) предназначен для обработки больших данных (внешнего сетевого потока данных) и генерации специализированных правил фильтрации. В частности, используется при возникновении ситуации Δ_4 (контур управления, раздел 3.1), требующей подключения эксперта #2 в двух случаях: при существенном изменении сетевого потока данных (внешнее возмущающее воздействие); при некорректном формировании правила обработки информации (фильтрации сетевого потока данных). Например, может возникнуть следующая ситуация: руководство предприятия обновили датчики и счетчик, в которых данные передаются по новому протоколу и с разным типом и видом сетевого потока данных.

Реализованная функциональность:

- считывание скорости сетевого потока данных с внешнего сетевого интерфейса;
- добавление правил фильтрации через веб-часть;
- анализ и фильтрация входящего/исходящего сетевого потока данных на основе импульсной нейронной сети;
- вывод скорости внешнего сетевого интерфейса и количества сетевых пакетов в веб-часть;
- автоматизированный (поминутный) анализ скорости и количества сетевых пакетов с возможностью изменения лимитов в веб-части для определения наличия аномалий в сетевом потоке данных;
- сохранение данных в СУБД MySQL с дальнейшим выводом в веб-интерфейс;
- блокировка IP-адресов через веб-интерфейс.

Блокировка определенного IP-адреса представлена на рисунке 3.15.



Блокировка IP-адресов	
IP-адрес	
Время блокировки	

Рисунок 3.15 – Блокировка адреса через веб-интерфейс на основе принятого решения экспертом согласно реализованному ранее контуру управления

Основная функциональность представлена в таблице 3.1.

Таблица 3.1 – Описание основных выполняемых команд

Название	Цель выполнения	Условие выполнения	Полезность при выполнении
Старт	Запуск модуля: автоматический и ручной (из веб-части и в консольном режиме)	Работа физического сервера	Анализ входящего и исходящего сетевого потока данных. Автоматическое применение правил при наличии большого сетевого потока данных
Стоп	Остановка модуля: автоматическая и ручная (из веб-части и в консольном режиме)	Ручное выполнение команды	Уведомление о сетевом состоянии системы в веб-интерфейс. Деактивация модуля
Перезапуск	Перезагрузка модуля из веб-части и консоли	Ручное выполнение команды	Меньшее потребление ресурсов
Проверка входящего и исходящего сетевого потока данных	Выявление наличия аномалий в сетевом потоке данных и применение правил фильтрации	Работа физического сервера	Фильтрация внешнего сетевого потока данных на основе нейронной сети и IPTables
Мониторинг датчиков электроэнергетического оборудования	Снятие показаний для обучения ИМНС и РНС	Работа физического сервера и автоматическое выполнение команды	Формирование нового датасета для переобучения нейронных сетей
Установка лимита скорости и входящих сетевых пакетов через веб-часть	Задать собственное значение лимита входящего и исходящего сетевого потока данных, сетевых пакетов	Достижение лимитированного значения, заданного из веб-части	Реализовывает лимитированный порог сетевого потока данных в виде сетевых пакетов

Правило фильтрации внешнего сетевого потока данных через IPTables представляет собой последовательное выполнение команд, преобразованных в одну единую цепочку с задаваемыми параметрами. Примеры некоторых реализованных правил представлены в таблице 3.2. Данные правила определяют аномалии сетевого потока данных по количеству входящих сетевых пакетов, а также позволяют его отфильтровать. Соответственно, правила фильтрации автоматически прописываются нейронной сетью в IPTables.

Таблица 3.2 – Реализация правил фильтрации с предполагаемой нагрузкой

Название	Применение	Тип протокола	Правило	Предполагаемая нагрузка на CPU, %
Фильтр HTTP	Фильтрация 80 порта	HTTP	<code>iptables -A INPUT -p udp -m tcp -dport 80 -m connlimit --connlimit-above 1000 -j DROP</code>	5,00
Фильтр HTTPS	Фильтрация 443 порта	HTTPS	<code>iptables -A INPUT -p udp -m tcp -dport 443 -m connlimit --connlimit-above 1000 -j DROP</code>	5,00
Фильтр SYN-пакетов	Фильтрация TCP	TCP	<code>iptables -A syn_flood -m limit --limit 500/s --limit-burst 2000 -j RETURN</code>	6,00
Фильтр сетевого потока данных по протоколу UDP	Фильтрация UDP	UDP	<code>iptables -A INPUT -p udp -m udp -dport 0:65535 -m state --state NEW -m recent --update --seconds 10 --hitcount 100 --name DEFAULT --resource -j DROP</code>	5,50

При этом у предлагаемых метода и модели существует ряд ограничений касаясь области применения: метод обработки больших данных для формирования

исходного датасета не привязан на прямую к геотехническому объекту, но модель нейронной сети структурно привязана к параметрам и количеству данных, которые необходимо фильтровать для формирования датасета. В отличие от стандартных методов, приведенных в первой главе диссертации, предложенный метод обладает большей универсальностью, что позволяет расширить его сферы применения. Однако, модель нейронной сети, содержащая структуру, архитектуру, конфигурацию и метод обучения подлежит подстройке для конкретных условий работы геотехнического объекта. В представленном контуре управления (раздел 3.1) реализуется возможность автоматизированного определения необходимости переобучения нейронной сети за счет анализа видов и типов информации (сетевой поток данных). Более того, корректное использование ИмНС возможно только при достаточно большом сформированном датасете и при неизменной конфигурации оборудования. При любом изменении конфигурации оборудования, либо переноса системы на другой объект, требуется переобучение ИмНС.

3.3.2. Основные возможности и ограничения метода и модели прогнозирования многомерных временных рядов на основе РНС в рамках прототипа СППР

Разработанный метод (РНС) предназначен для обработки и анализа больших темпоральных данных, а также для прогнозирования временных рядов. В частности, используется при возникновении ситуации $\Delta 3$, требующей подключения эксперта #1 (рисунок 3.2): случай, когда требуются прогнозные значения потребления электроэнергии для принятия управленческих решений, при этом параметры ИНС настроены некорректно, как следствие, меньшая точность

значений на выходе. Например, у руководства электроэнергетического предприятия появилась необходимость в резервировании мощностей для физических и юридических лиц на основе прогнозных значений потребления электроэнергии за счет данных за предыдущий отчетный период.

Реализованная функциональность:

– просмотр количества потребляемой электроэнергии по датам в виде графика. Позволяет по датам визуализировать представленную информацию в веб-интерфейсе (рисунок 3.16);

– прогноз временных рядов с использованием РНС и с последующим выводом результатов в веб-интерфейс в двух форматах: вещественный и целочисленный. Позволяет спрогнозировать значения потребляемой электроэнергии (рисунок 3.17) на основе модели искусственной рекуррентной нейронной сети, сохранить их в базу данных и вывести прогноз (рисунок 3.18) в веб-часть;

– обучение РНС из веб-интерфейса. Позволяет переобучать нейронную сеть нажатием одной кнопкой в случае некорректной подстройки весовых синаптических коэффициентов (рисунок 3.19);

– формирование исходного датасета в ручном режиме (рисунок 3.20). Позволяет в автоматическом режиме получать данные с внешних и внутренних источников для формирования исходного датасета в MySQL (рисунок 3.21) при помощи ИмНС с целью обучения РНС;

– формирование отчетов в формате .xlsx для проведения сравнительного анализа. Позволяет анализировать ошибку кВт*ч по всем электросчетчикам в целом, и по каждому в частности (рисунок 3.22);

– вывод карты прогнозов с расположением электросчетчиков. Позволяет просматривать географическое расположение счетчиков с их выводом в веб-интерфейс (рисунок 3.23).

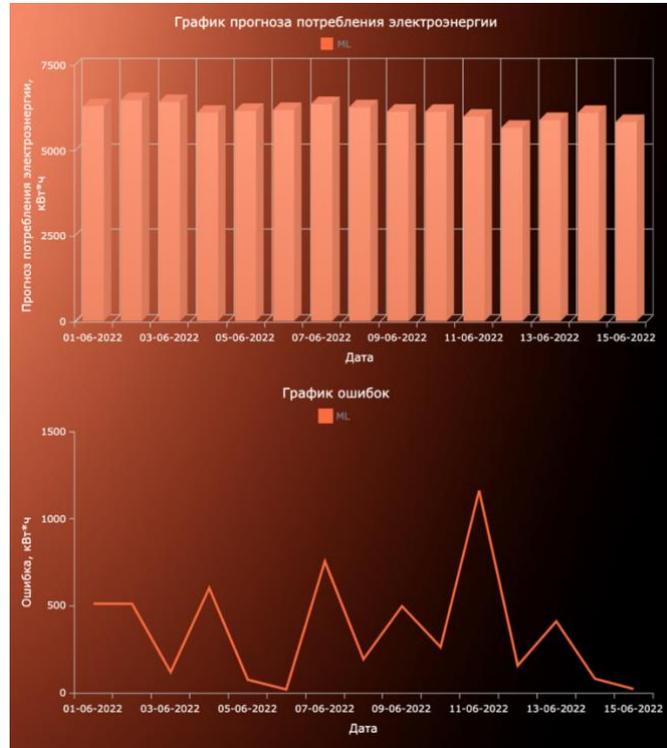


Рисунок 3.16 – Визуализированный вывод количества потребляемой электроэнергии и прогнозных ошибок

Выберите на основе каких столбцов будет произведено обучение

- Температура воздуха
- Атмосферное давление
- Выходной день или нет
- Сезон (зима, весна, лето, осень)
- Количество осадков

Выберите на основе каких моделей будет произведено обучение

ML, Java Neural Network, LSTM, ▼

Отправить

Рисунок 3.17 – Функция выполнения прогноза

Дата	Прогноз	Реальные значения	Ошибка
01-06-2022	6309.4766	5797.8	511.67676
02-06-2022	6487.931	6998.8	510.86865
03-06-2022	6427.773	6546.4	118.62695
04-06-2022	6116.822	6717.8	600.978
05-06-2022	6170.4463	6095.0	75.44629
06-06-2022	6196.4336	6177.0	19.433594
07-06-2022	6368.0947	5613.8	754.2949
08-06-2022	6277.273	6084.0	193.27295
09-06-2022	6149.982	5653.0	496.98193

Рисунок 3.18 – Вывод прогнозных значений из базы данных

Train

Оптимальные значения lr - в диапазоне от 0.7 до 0.95 (в зависимости от количества итераций)

Установите значение learningRate

0.9

Установите количество скрытых слоёв

1

Установите количество нейронов в скрытых слоях

9

Установите количество итераций

1500

Установите размер тренировочной выборки

0.75

Установите размер валидационной выборки

0.15

Установите размер тестовой выборки

0.1

Выберите количество дней на основе которых будет произведен запрос

1

Выберите количество дней для прогноза

1

Выберите на основе каких столбцов будет произведено обучение

- Температура воздуха
- Атмосферное давление
- Выходной день или нет
- Сезон (зима, весна, лето, осень)
- Количество осадков

Рисунок 3.19 – Функции обучения РНС и резервной копии моделей нейронной сети

Парсер

Получить данные из внешних и внутренних источников

Рисунок 3.20 – Функция получения данных из внешних источников с последующим внесением в MySQL

- elforecast
- Новая
- complete_with_electro
- forecast1
- forecast2
- forecast3
- Новая
- complete
- complete_final
- complete_test
- complete_with_electro
- data
- full_result_complete
- full_result_complete_with_electro
- full_result_complete_xml
- full_result_complete_xml1
- xmlparsertest
- information_schema
- performance_schema

✔ Отображение строк 0 - 24 (1277 всего, Запрос занял 0.0007 сек.)

SELECT * FROM `complete_with_electro`

Профилирование [[Построчное редактирование](#)] [[Изменить](#)] [[Анализ SQL запроса](#)] [[Создать PHP-код](#)] [[Обновить](#)]

> >> | Количество строк: | Фильтровать строки:

+ Параметры							
Дата	№ счетчика	Потребляемая электроэнергия, кВт*ч	Температура воздуха	Количество осадков	Выходной день или нет	Атмосферное давление	Сезон (зима, весна, лето, осень)
2018-01-01	1	7156,5	12	0	0	1022	1
2018-01-02	1	7567,1	14	0	0	1024	1
2018-01-03	1	8123,2	15	0	0	1021	1
2018-01-04	1	8119,3	15	0	0	1022	1
2018-01-05	1	8912	16	0	0	1021	1
2018-01-06	1	7118,4	15	0	1	1022	1
2018-01-07	1	7945,2	16	0	1	1022	1
2018-01-08	1	7981,5	16	0	0	1018	1
2018-01-09	1	8589	15	0	0	1021	1
2018-01-10	1	8611,7	17	0	0	1018	1

Рисунок 3.21 – Сформированный для прогнозирования датасет, хранимый в MySQL

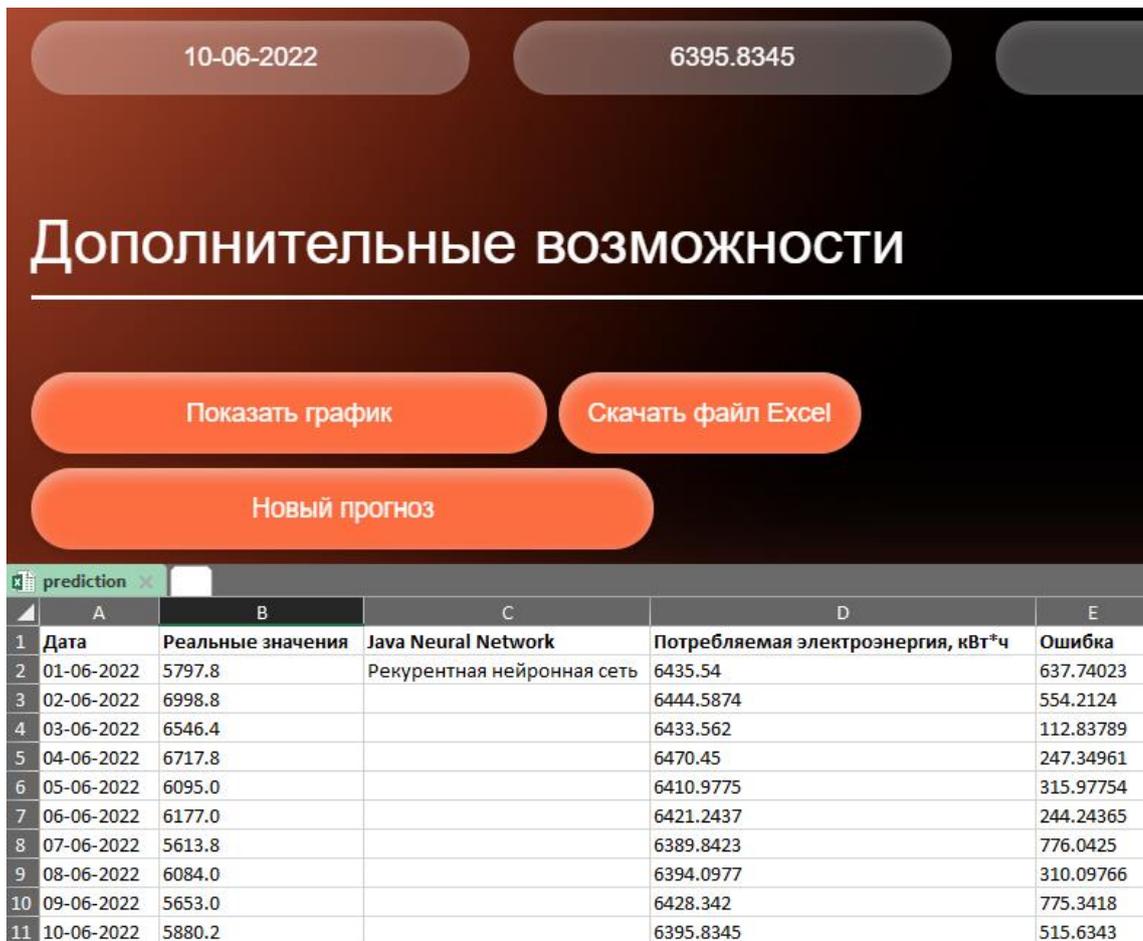


Рисунок 3.22 – Функция формирования отчета о прогнозных значениях

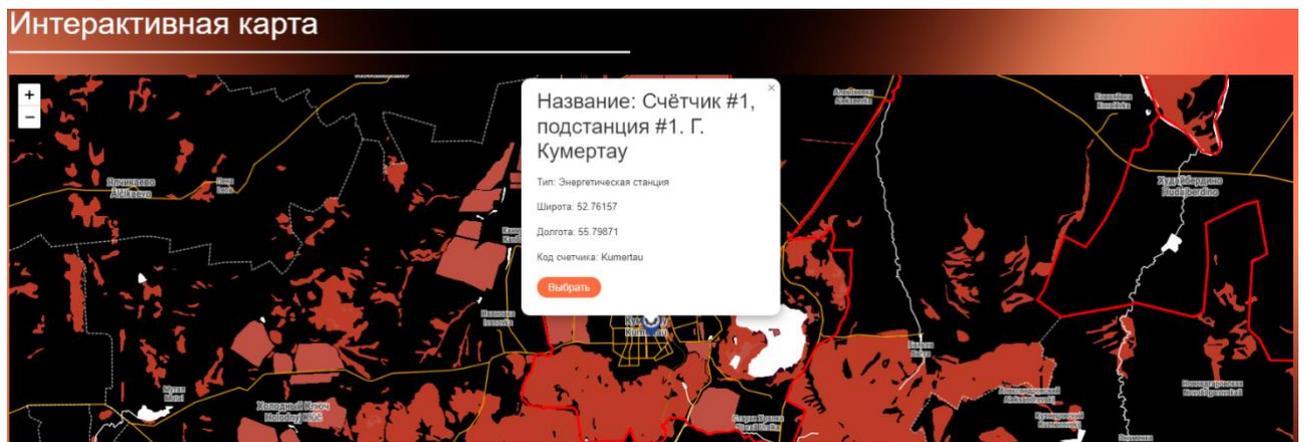


Рисунок 3.23 – Интерактивная карта с географическим расположением счетчиков

Основные ограничения. Во-первых, ограничения касаются некоторых прогнозных всплесков (например, при прогнозировании на 6-й день, исходя из результатов эксперимента таблица 3). Это связано с особенностями

фундаментально доработанного нейрона и метода обучения. В частности, есть предположение о том, что в некоторых случаях коэффициенты подбираются с небольшой погрешностью (ошибкой). Для этого также необходимы дополнительные исследования.

Во-вторых, ограничения касаются сравнения с методами обучения Machine Learning. В связи с тем, что исходные данные, подаваемые на глубокое обучение в предлагаемую ИНС, являются разнородными, что не позволяет добиться высокой точности при решении задач с использованием регрессии. Deep Learning, в отличие от Machine Learning, позволяет достигать достаточно высокой точности выходных значений по причине наличия слоев в ИНС. В будущих исследованиях планируется разработать методы анализа и восстановления данных для решения данной проблемы, что позволит проводить дополнительный сравнительный анализ относительно Machine Learning.

Выводы по главе 3

1. Разработан прототип СППР, основанный на нейросетевой обработке больших темпоральных данных и прогнозировании временных рядов. Разработанное программное обеспечение позволяет свести к минимуму риски финансовых убытков предприятия за счет принятия решений ЛПР и обладает следующей основной функциональностью: информационная поддержка принятия управленческих решений на основе автоматизированной обработки больших разнородных и темпоральных данных; прогнозирование временных рядов и генерация правил фильтрации сетевого потока данных на основе автоматизированной обработки больших данных, а также методов и моделей, представленных в главе 2. Прототип состоит из объединенных модулей (ИМНС, РНС, RCM и веб-часть) и реализован на следующем стеке технологий:

a. Backend: языки программирования «Java» и «Python»; «Apache Maven» + «Spring Boot»; ORM (Object-Relational Mapping) + JPA (Java Persistence API) + Hibernate; СУБД «MySQL».

b. Frontend: HTML, CSS, Bootstrap, OpenStreetMap, JavaScript.

2. Разработан четырехкаскадный контур управления в виде четырех контуров управления, в рамках которого выделен ОУ – информационная поддержка принятия управленческих решений. Управленческие решения могут быть связаны с: резервированием дополнительных мощностей электроэнергетического предприятия на основе прогнозных значений потребления электроэнергии, в том числе и с проведением энергосберегающих мероприятий; внедрением правил фильтрации сетевого потока данных для решения задач оптимизации сети различных организаций, в том числе и энергетических предприятий, а также центров обработки данных.

3. Описана общая функциональная структурная схема в рамках процесса автоматизации разработанного прототипа СППР:

- обеспечение автоматического взаимодействия между каждым модулем и веб-частью;
- реализация платформы для бесконфликтной и взаимосвязанной работы модулей между собой;
- создание связей между базой данных MySQL и модулями для дальнейшего хранения передаваемой информации;
- обработка и фильтрация больших объемов сетевого потока данных, а также прогнозирование временных рядов для поддержки принятия решений в области электроэнергетики.

4. Описана основная функциональность прототипа СППР по обработке больших разнородных и темпоральных данных и прогнозированию временных рядов. Метод, представленный в виде модели ИмНС (модуль): считывание скорости сетевого потока данных с внешнего сетевого интерфейса; добавление правил фильтрации через веб-часть; анализ и фильтрация входящего/исходящего сетевого потока данных на основе импульсной нейронной сети; вывод скорости внешнего сетевого интерфейса и количества сетевых пакетов в веб-часть; автоматизированный (поминутный) анализ скорости и количества сетевых пакетов с возможностью изменения лимитов в веб-части для определения наличия аномалий в сетевом потоке данных; сохранение данных в СУБД MySQL с дальнейшим выводом в веб-интерфейс; блокировка IP-адресов через веб-интерфейс.

Метод, представленный в виде модели РНС (модуль):

- просмотр количества потребляемой электроэнергии по датам в виде графика. Позволяет по датам визуализировать представленную информацию в веб-интерфейсе;
- прогноз временных рядов с использованием РНС с последующим выводом результатов в веб-интерфейс в двух форматах: вещественный и целочисленный. Позволяет на основе модели искусственной рекуррентной нейронной сети

спрогнозировать значение потребляемой электроэнергии, сохранить в его в базу данных и вывести прогноз в веб-часть;

– обучение РНС из веб-интерфейса. Позволяет переобучать нейронную сеть нажатием одной кнопки в случае некорректной подстройки весовых синаптических коэффициентов;

– бекап (резервная копия) моделей РНС. Позволяет создавать резервные копии моделей ИНС на текущую дату при некорректной подборке весовых синаптических коэффициентов;

– внесение данных в СУБД MySQL. Позволяет заносить данные через веб-интерфейс в базу данных с целью переобучения нейронной сети с учетом новых данных;

– формирование отчетов в формате .csv для проведения сравнительного анализа. Позволяет анализировать ошибку кВт*ч по всем электросчетчикам в целом, и по каждому в частности;

– вывод карты прогнозов с расположением электросчетчиков. Позволяет просматривать географическое расположение счетчиков с их выводом в веб-интерфейс.

Посредством реализованной функциональности принимались различные управленческие решения: например, резервирование мощностей на основе прогнозных значений потребляемой электроэнергии; ограничение доступа к конфиденциальным данным на основе сгенерированных правил обработки информации (фильтрации сетевого потока данных).

ГЛАВА 4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРЕДЛОЖЕННЫХ РЕШЕНИЙ НА ПРИМЕРЕ ПРЕДПРИЯТИЯ ЭНЕРГЕТИЧЕСКОГО СЕКТОРА

Оценка эффективности предложенных решений (обработка больших объемов сетевого потока данных и прогнозирование потребления электроэнергии) производилась по следующим критериям: качество (Data Quality – точность и своевременность) обработки данных за счет сгенерированных правил фильтрации с их последующей нагрузкой на ресурсы кластера, точность прогнозных значений и скорость их расчета за счет обучения модифицированным методом обратного распространения ошибки. Учитывая, что предложенные методы и модели применены при разработке прототипа СППР, эффективность предложенных решений будем оценивать на прототипе СППР.

Хотелось бы отметить, что все типы СППР подразделяются на [178, 179]:

1. СППР, *использующие независимые витрины данных*. Применяются в крупных организациях, имеющих несколько подразделений, в том числе и отделы информационных технологий.

2. СППР *на основе двухуровневого хранилища данных*. Используются в крупных компаниях, данные которых консолидированы в единую систему.

3. СППР *на основе трехуровневого хранилища данных*. В них используются хранилища данных, из которых формируются витрины данных для решения группами пользователей схожих (однотипных) задач.

4. *Функциональные СППР*, к которым и относится разработанный прототип СППР. Используются в организациях, в которых важно использовать одну платформу (например, операционную систему) и оперативность в связи с отсутствием необходимости перегружать данные в специализированное хранилище.

При этом основным недостатком при использовании таких СППР является необходимость создания отдельных баз данных для решения той или иной задачи,

что существенно повышает нагрузку на ресурсы физического сервера из-за огромного количества запросов к БД.

Так как наиболее критичным для функциональных типов СППР является повышение величины нагруженности вычислительных ресурсов, для каждого из рассмотренных моделей и методов дополнительно будем рассматривать этот показатель. Исходя из анализа методов и моделей раздела 1.2 и требований из раздела 3.3, эффективность предложенных решений прототипа СППР измеряется в следующих критериях:

1. Скорость обработки данных за счет сгенерированных правил фильтрации с их последующей нагрузкой на ресурсы физического сервера (обучение ИмНС и формирование датасета)
2. Прогнозная точность и применение в электроэнергетике (показатель относительной/предполагаемой эффективности).
3. Ресурсоемкость процесса обработки данных и прогнозирования временных рядов.

При этом приводятся результаты экспериментальных исследований, демонстрирующих, что использование разработанного прототипа системы поддержки принятия решений является целесообразным и эффективным при обработке больших разнородных и темпоральных данных на примере обработки больших объемов данных и прогнозирования значений потребления электроэнергии, что также позволило автоматизировать процесс прогнозирования и обработки данных для предприятий энергетического сектора.

В связи с тем, что прототип СППР представлен несколькими программными модулями, реализующими предложенные методы и модели, необходима оценка отдельных модулей по вышеуказанным критериям, а оценку прототипа СППР оценить именно по ресурсоемкости (величиной нагрузки на физические ресурсы вычислительной сети).

Таким образом, главу 4 можно разделить на 3 составляющих:

1. Оценка импульсной нейронной сети (метод №1, раздел 2.2) – скорость обработки данных (с учетом времени обучения) за счет сгенерированных правил фильтрации и нагрузка на вычислительные ресурсы.

2. Оценка прогнозной точности рекуррентной нейронной сети (метод №2, раздел 2.3).

3. Оценка эффективности отдельных модулей прототипа СППР и совокупная нагрузка прототипа СППР на физические ресурсы.

4.1. Оценка эффективности метода и модели обработки сетевого потока данных с применением импульсной нейронной сети на примере предприятия электроэнергетического сектора

В данном разделе представлено сравнение (качественных и количественных параметров) методов самообучения ИНС и апробация ИмНС при фильтрации больших объемов сетевого потока данных.

4.1.1. Апробация ИмНС

В данном пункте представлены результаты: апробации разработанной системы; сравнительного анализа функциональных возможностей относительно аналогичных решений; оценки производительности, а также скорости и точности обучения ИмНС в сравнении с другими методами. В качестве инструментов анализа сетевого потока данных выступают два решения: предлагаемый прототип СППР на вычислительном кластере и DPI (Deep Packet Inspection, «Hippie») на серверах, находящихся в сети CDN.

Для выявления эффективности предлагаемого метода, реализованного в виде модели ИмНС был проведен анализ функциональных возможностей аналогичных методов, результаты которого представлены в таблице 4.1: стандартные решения (например, IPTables – NetFilter); DPI – (Cisco, Sandvine, Huawei). При этом модель

импульсной нейронной сети (ИмНС) подпадает под ситуации, описываемые в рамках контура управления и системной модели (раздел 3.1):

а) Δ_1 – ситуация, не требующая вмешательства ЛПР: случай, когда выходные данные удовлетворяют ЛПР/экспертов.

б) Δ_2 – ситуация, требующая вмешательства ЛПР, но не требующая подключения эксперта #1: случай, когда требуются необходимые правила фильтрации или прогнозные значения для принятия управленческих решений.

в) Δ_4 – ситуация, требующая подключения эксперта #2 в двух случаях: при существенном изменении сетевого потока данных (внешнее возмущающее воздействие); при некорректном формировании правила фильтрации сетевого потока данных.

Таблица 4.1 – Сравнение функциональных возможностей ИмНС и других методов, в том числе и программных решений

Общая функциональность и возможности	Приоритет	ИмНС	Стандартные решения	Решения DPI
Работа с большими объемами данных	1	Да	Нет	Да
Скоростное обучение	2	Да	Нет	Нет
Реагирование на внешние ситуации	3	Да	Нет	Да
Возможность автоматического выявления аномалий в сетевом потоке данных	4	Да	Нет	Нет
Высокое время реагирования (до 20 сек.)	5	Да	Нет	Нет
Повышение задержки	6	Нет	Да	Да
Сохранение новых правил обработки информации в базу данных	7	Да	Нет	Нет
Подробный анализ сетевых пакетов	8	Да	Нет	Да
Большое потребление серверных ресурсов при большом количестве разнородных данных	9	Нет	Да	Да

Таким образом, исходя из таблицы 4.1, необходимо отметить, что предлагаемый метод в виде импульсной нейронной сети не только более оптимизирован под высокие нагрузки и способен обрабатывать большие объемы данных, но и в автоматическом режиме обнаруживает аномалии в сетевом потоке данных с его дальнейшей фильтрацией. Это позволяет не только эффективно идентифицировать какие-либо ситуации, связанные с обработкой больших объемов данных, но и формировать целые датасеты в виде темпоральной информации с различных датчиков и счетчиков для прогнозирования реализованной рекуррентной нейронной сетью.

Далее в рамках реализации прототипа СППР была произведена апробация разработанной импульсной нейронной сети в обычном режиме — без большого количества поступающих данных (таблица 4.2) и при большом количестве поступающих разнородных данных – активный режим (таблица 4.3). Изначальная установка перед экспериментом: емкость канала 20 Гбит/сек, при этом 10 Гбит/сек активно используются. Таким образом, сможем максимально задействовать только 9.9 Гбит/сек, оставив при этом 100 Мбит/сек для снятия нагрузочных показаний.

Таблица 4.2 – Потребление ресурсов в обычном режиме

Ситуация	День эксперимента (p_i)									
	1	2	3	4	5	6	7	8	9	10
	Скорость потока данных $N_c^h(t)$, Гбит/сек									
	1,0	1,2	1,8	2,0	2,5	2,6	2,7	2,8	2,9	3,5
	Загруженность CPU (RND_{cpu}^{sc}), %									
Старт (st_i)	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,3
Перезагрузка (res_i)	0,2	0,3	0,2	0,1	0,2	0,2	0,3	0,3	0,3	0,4
Анализ сетевого потока данных (an_i)	0,4	0,4	0,5	0,5	0,5	0,6	0,7	0,7	0,8	0,9
	Загруженность SSD (RND_{ssd}^{sc}), %									
Старт (st_i)	0,1									
Перезагрузка (res_i)	0,1	0,2	0,2	0,3	0,3	0,4	0,5	0,6	0,7	0,8
Анализ сетевого потока данных (an_i)	0,3	0,4	0,6	0,7	0,8	1,0	1,1	1,2	1,4	1,5
	Загруженность ОЗУ (RND_{ram}^{sc}), %									
Старт (st_i)	0,1									
Перезагрузка (res_i)	0,2	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
Анализ сетевого потока данных (an_i)	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2	1,3	1,4

Таблица 4.3 – Потребление ресурсов в активном режиме

Ситуация	День эксперимента (p_i)									
	1	2	3	4	5	6	7	8	9	10
	Скорость сетевого потока данных ($N_c^h(t)$), Гбит/сек									
	2,3	3,5	4,0	4,4	4,8	5,8	6,4	7,9	8,3	9,9
	N_{np} , млн. шт./сек									
	0,5	1,0	1,5	2,0	2,5	3,0	4,5	7,0	9,5	9,9
	Загруженность CPU ($RD_{cpuload}^{sc}$), %									
Старт (st_i)	3,1	4,2	4,6	4,9	5,6	6,0	7,5	8,0	9,0	9,3
Перезагрузка (res_i)	4,1	4,8	5,2	6,3	7,0	7,5	8,7	9,0	9,5	9,7
Анализ сетевого потока данных (an_i)	5,2	5,6	6,0	6,5	7,3	8,2	8,6	8,9	9,3	9,9
	Загруженность SSD ($RD_{ssdload}^{sc}$), %									
Старт (st_i)	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2,3
Перезагрузка (res_i)	2,0	3,2	4,1	4,3	4,8	5,0	5,5	6,0	6,7	7,5
Анализ сетевого потока данных (an_i)	2,4	3,0	3,5	4,0	4,7	5,6	7,0	8,2	8,6	9,3
	Загруженность ОЗУ ($RD_{ramload}^{sc}$), %									
Старт (st_i)	0,8	0,9	1,3	1,5	1,9	2,3	2,8	3,0	3,2	3,7
Перезагрузка (res_i)	1,3	1,6	1,9	2,5	2,9	3,8	4,0	5,1	5,7	5,9
Анализ сетевого потока данных (an_i)	1,7	1,9	2,3	2,7	3,6	4,0	4,5	5,2	5,7	6,4

Таким образом, по данным таблиц 4.2–4.3 была рассчитана на основании формул (4.1–4.6) средняя нагрузка по загруженности CPU, SSD и ОЗУ, что представлено в таблице 4.4.

Таблица 4.4 – Средняя нагрузка на серверные ресурсы

Ресурс	Обычный режим	Активный режим
Центральный процессор, %	0,3	6,9
Твердотельный накопитель, %	0,4	4,0
Оперативная память, %	0,5	3,1

$$RND_{cpuload}^{sc} = \frac{\sum_{i=1}^p RND_{cpuload}^{sc}(i)}{\sum_{i=1}^p (st_i + res_i + an_i)}, \quad (4.1)$$

$$RNDS_{ssdload}^{sc} = \frac{\sum_{i=1}^p RND_{ssdload}^{sc}(i)}{\sum_{i=1}^p (st_i + res_i + an_i)}, \quad (4.2)$$

$$RNDS_{ramload}^{sc} = \frac{\sum_{i=1}^p RND_{ramload}^{sc}(i)}{\sum_{i=1}^p (st_i + res_i + an_i)}, \quad (4.3)$$

$$RDS_{cpuload}^{sc} = \frac{\sum_{i=1}^p RD_{cpuload}^{sc}(i)}{\sum_{i=1}^p (st_i + res_i + an_i)}, \quad (4.4)$$

$$RDS_{ssdload}^{sc} = \frac{\sum_{i=1}^p RD_{ssdload}^{sc}(i)}{\sum_{i=1}^p (st_i + res_i + an_i)}, \quad (4.5)$$

$$RDS_{ramload}^{sc} = \frac{\sum_{i=1}^p RD_{ramload}^{sc}(i)}{\sum_{i=1}^p (st_i + res_i + an_i)}, \quad (4.6)$$

в которых $RNDS_{cpuload}^{sc}$ – средняя загруженность CPU в режиме простоя, $RNDS_{ssdload}^{sc}$ – средняя загруженность SSD в режиме простоя, $RNDS_{ramload}^{sc}$ – средняя загруженность RAM в режиме простоя, $RDS_{cpuload}^{sc}$ – средняя загруженность CPU в активном режиме, $RDS_{ssdload}^{sc}$ – средняя загруженность SSD в активном режиме, $RDS_{ramload}^{sc}$ – средняя загруженность RAM в активном режиме, p – количество дней эксперимента.

Нагрузка на ресурсы всех физических серверов кластера достаточно низкая, несмотря на массивные потоки больших разнородных данных. Все физические серверы во время поступления больших объемов сетевого потока данных были доступны по внешней глобальной сети и работали без перебоев за счет быстрого

обучения и реагирования на аномалии во внешнем сетевом потоке данных разработанной импульсной нейронной сетью.

Сравнение методов обучения импульсной нейронной сети представлено в таблице 4.5.

Таблица 4.5 – Сравнение методов обучения ИмНС без учителя

Параметр	STDP	GUL	ВСМ- правило	ABS- правило	Предложенный метод обучения
Скорость обучения «с нуля» при больших объемах сетевого потока данных, мин	70,00	60,00	150,00	80,00	20,00
Максимально возможное количество обрабатываемых данных перед кодированием в импульсы ⁶ , Гбит/с	10,00	12,00	11,65	10,54	13,19
Возможность обучения нейронной сети с расчетом текущей занятости вычислительных ресурсов	НЕТ				ДА

Таким образом, предложенный метод обучения показал:

- достаточно высокую скорость (20 минут на полное обучение) при самообучении для фильтрации больших объемов внешнего сетевого потока данных;

- возможность обработки большего объема данных (13,19 Гбит/с), нежели при остальных рассматриваемых методах обучения;

- возможность расчета вычислительных ресурсов кластера для самообучения импульсной нейронной сети. Это позволяет не перегружать ресурсы кластера и параллельно запускать другие ресурсоемкие процессы, необходимые для решения тех или иных задач.

⁶ Зависит от метода обучения и оборудования: чем оно мощнее, тем больше можно обработать данных

Тестирование проводилось на следующем оборудовании:

- количество физических серверов — 30;
- процессор Intel Xeon 5690 (CPU — 60, физических ядер — 360, количество потоков — 720);
- оперативная память — 960 ГБ;
- SSD — RAID 10 (NVME Samsung 980 pro 500ГБ каждый);
- внешний сетевой канал — 20 Гбит/с;
- внутренний сетевой канал (локальная сеть) — 100Гбит/с.

Данный кластер соответствует требованиям текущего диссертационного исследования в рамках обработки большого объема информации, в том числе вычисления прогнозных значений RNN и анализа потоковых данных.

Также было принято решение сравнить нагрузку при обработке больших разнородных данных с реализованной для специализированного интернет-ресурса электроэнергетического предприятия сетью CDN, архитектура которой представлена на рисунке 4.1.

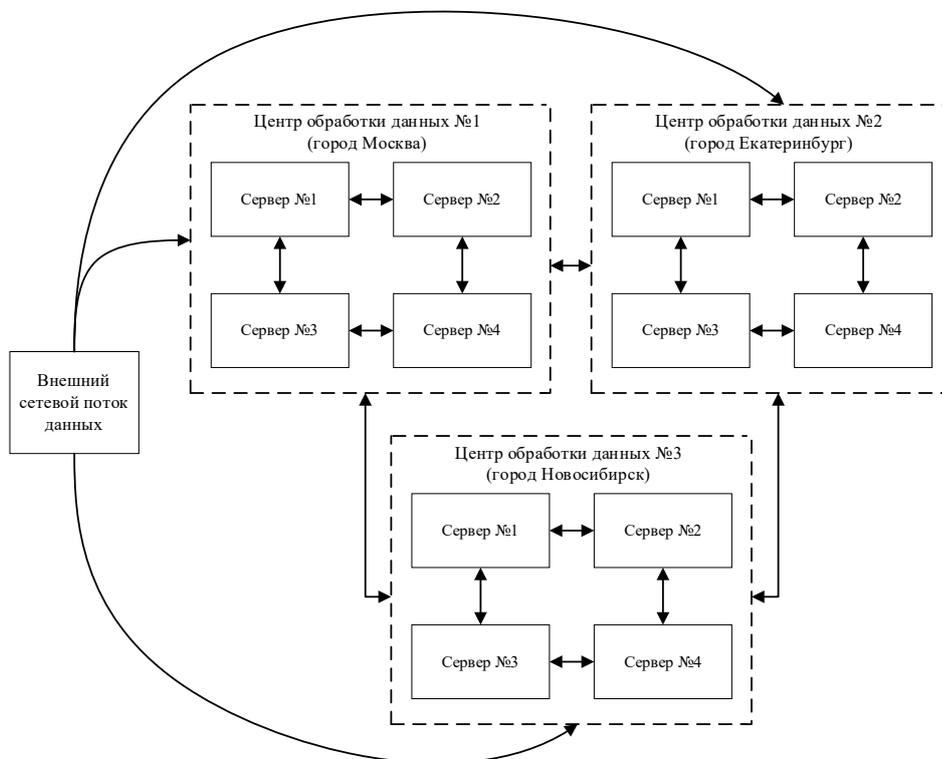


Рисунок 4.1 – Реализованная сеть CDN

Такая модель сети может использоваться в электроэнергетике в случае удаленного объединения нескольких систем в единое информационное пространство.

У данной сети CDN также реализована панель управления, в которой отображается статистика не только кэшированного и не кэшированного внешнего сетевого потока данных, но и нагрузка на каждый физический сервер по отдельности и на все серверы в сумме. Для чистоты проведения эксперимента (результаты представлены в таблицах 4.6–4.7) используемый объем сетевого потока данных и количество сетевых пакетов, а также внешний сетевой канал полностью идентичны эксперименту, результаты и данные которого представлены в таблицах 4.2–4.5. В таблице 4.6 $N_{cdn}^h(t)$ – сетевой, аналогичный $N_c^h(t)$ поток данных в сети CDN.

Таблица 4.6 – Потребление ресурсов в обычном режиме (CDN)

Ситуация	День эксперимента (p_i)									
	1	2	3	4	5	6	7	8	9	10
	Скорость сетевого потока данных $N_{cdn}^h(t)$, Гбит/сек									
	1,0	1,2	1,8	2,0	2,5	2,6	2,7	2,8	2,9	3,5
	Загруженность CPU ($CDND_{cpu}^{sc}$), %									
Старт (st_i)	0,20	0,22	0,24	0,30	0,29	0,35	0,38	0,42	0,45	0,49
Перезагрузка (res_i)	0,40	0,40	0,50	0,50	0,40	0,50	0,50	0,50	0,50	0,6
Анализ сетевого потока данных (an_i)	0,8	0,80	0,80	0,80	0,80	0,90	0,90	0,90	0,93	0,15
	Загруженность SSD ($CDND_{ssd}^{sc}$), %									
Старт (st_i)	0,25									
Перезагрузка (res_i)	0,28	0,29	0,35	0,38	0,40	0,42	0,45	0,68	0,76	0,99
Анализ сетевого потока данных (an_i)	0,30	0,40	0,60	0,70	0,80	1,00	1,10	1,20	1,40	1,50
	Загруженность ОЗУ ($CDND_{ram}^{sc}$), %									
Старт (st_i)	0,35									
Перезагрузка (res_i)	0,29	0,36	0,48	0,52	0,66	0,79	0,95	1,30	1,56	2,12
Анализ сетевого потока данных (an_i)	0,79	0,99	1,74	2,16	2,83	3,90	4,41	5,31	7,28	9,01

Таблица 4.7 – Потребление ресурсов в активном режиме (CDN)

Ситуация	День эксперимента (p_i)									
	1	2	3	4	5	6	7	8	9	10
	Скорость сетевого потока данных ($N_{cdn}^h(t)$), гбит/сек									
	2,3	3,5	4,0	4,4	4,8	5,8	6,4	7,9	8,3	9,9
	N_{np} , млн. шт./сек									
	0,5	1,0	1,5	2,0	2,5	3,0	4,5	7,0	9,5	9,9
	Загруженность CPU (CDN_{cpu}^{sc}), %									
Старт (st_i)	10,25	11,99	13,68	15,89	19,71	25,41	28,43	32,44	35,00	39,05
Перезагрузка (res_i)	15,10	18,19	23,34	29,00	34,12	37,18	40,12	42,16	45,50	49,00
Анализ сетевого потока данных (an_i)	17,20	23,78	26,90	35,44	38,30	44,11	47,54	60,12	64,88	67,53
	Загруженность SSD (CDN_{ssd}^{sc}), %									
Старт (st_i)	10,98	13,14	14,15	24,10	26,15	29,61	38,17	41,12	42,32	45,13
Перезагрузка (res_i)	16,14	21,12	28,44	33,19	37,18	47,00	51,01	55,00	59,05	60,00
Анализ сетевого потока данных (an_i)	20,41	28,47	31,85	35,78	44,97	49,53	58,71	68,14	72,69	79,83
	Загруженность ОЗУ (CDN_{ram}^{sc}), %									
Старт (st_i)	14,08	19,27	22,39	26,15	27,51	31,34	38,12	41,05	48,31	51,08
Перезагрузка (res_i)	18,31	24,60	26,00	29,00	34,12	45,10	47,00	49,00	61,00	61,82
Анализ сетевого потока данных (an_i)	24,71	24,99	31,80	34,00	39,12	46,13	56,17	64,22	74,10	79,00

Таким образом, по данным таблиц 4.6–4.7 по аналогии с формулами (4.1–4.6) были рассчитаны средние значения загрузки CPU, SSD и ОЗУ, что представлено в таблице 4.18.

Таблица 4.8 – Средняя нагрузка на ресурсы в сети CDN

Ресурс	Обычный режим	Активный режим
Центральный процессор, %	0,53	33,04
Твердотельный накопитель, %	0,67	39,44
Оперативная память, %	2,27	39,64

Анализ удельного веса снижения нагрузки на вычислительный кластер импульсной нейронной сетью (ИмНС) и CDN показан в таблице 4.9. Расчеты проводились по следующей системе уравнений:

$$\begin{cases} UD1 = \frac{(\text{Обычный режим}(CDN) - \text{Обычный режим}(ИмНС)) \cdot 100}{\max(\text{Обычный режим}(CDN), \text{Обычный режим}(ИмНС))} \\ UD2 = \frac{(\text{Активный режим}(CDN) - \text{Активный режим}(ИмНС)) \cdot 100}{\max(\text{Активный режим}(CDN), \text{Активный режим}(ИмНС))} \end{cases}, \quad (4.7)$$

где $UD1$ и $UD2$ – удельные веса снижения нагрузки при обычном и активном режимах.

Таблица 4.9 – Средняя нагрузка на ресурсы в сети CDN и вычислительный кластер ИмНС

Ресурс	Обычный режим (ИмНС)	Обычный режим (CDN)	$UD1$, %	Активный режим (ИмНС)	Активный режим (CDN)	$UD2$, %
Центральный процессор, %	0,30	0,53	43,39	6,90	33,04	79,11
Твердотельный накопитель, %	0,40	0,67	40,20	4,00	39,44	89,85
Оперативная память, %	0,50	2,27	77,97	3,10	39,64	92,17

Таким образом, удельные веса снижения нагрузки вычислительного кластера ИмНС и CDN следующие:

– *обычный режим.* На центральном процессоре *UD1* составляет 43,39% в пользу вычислительного кластера ИмНС, так как нагрузка меньше на вычислительном кластере;

– *обычный режим.* На SSD *UD1* составляет 40,20% в пользу вычислительного кластера ИмНС, так как нагрузка меньше на вычислительном кластере;

– *обычный режим.* На RAM *UD1* составляет 77,97% в пользу вычислительного кластера ИмНС, так как нагрузка меньше на вычислительном кластере;

– *активный режим.* На центральном процессоре *UD2* составляет 79,11% в пользу вычислительного кластера ИмНС, так как нагрузка меньше на вычислительном кластере;

– *активный режим.* На SSD *UD2* составляет 89,85% в пользу вычислительного кластера ИмНС, так как нагрузка меньше на вычислительном кластере;

– *активный режим.* На RAM *UD2* составляет 92,17% в пользу вычислительного кластера ИмНС, так как нагрузка меньше на вычислительном кластере.

Оборудование сети CDN является следующим:

- количество физических серверов — 30;
- процессор Intel Xeon E5-2637v4 (CPU — 60, физических ядер — 240, количество потоков — 480);
- оперативная память — 960 ГБ;
- SSD — RAID 10 (NVME Samsung 980 pro 500ГБ каждый);
- внешний сетевой канал — 20 Гбит/с.

По мощности процессоры Intel Xeon 5690 и Intel Xeon E5-2637v4 являются идентичными, что представлено в рейтинге «CPU Mark» [149] от компании PassMark® Software Pty Ltd. Таким образом, данное оборудование фактически полностью идентично оборудованию вычислительного кластера,

предоставленного для проведения расчетов, в связи с чем можно считать эксперимент достоверным и объективным.

4.2. Оценка эффективности метода и модели прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети на примере предприятий энергетического сектора

Для проведения анализа эффективности предлагаемой интеллектуальной нейросетевой системы в рамках прогнозирования значений количества потребляемой электроэнергии использовались ретроспективные данные измерений за последние 4 года. Эти данные предоставлены производственным отделением «Кумертауские электрические сети» с 01.01.2018 г. по 21.06.2022 г. в следующем виде (база данных, таблица 4.10): дата, активная потребляемая электроэнергия, температура воздуха. Также были добавлены дополнительные данные из внешних источников, способствующих повышению прогнозной точности: количество осадков, выходной день и сезон. Название используемых счетчиков: «Меркурий 230 ART».

Согласно контуру управления (рисунок 3.2), все ситуации учтены в рамках проведения эксперимента:

1. Δ_1 – ситуация, не требующая вмешательства ЛПР: случай, когда выходные данные удовлетворяют ЛПР/экспертов. Например, выявлено отклонение, не требующее вмешательства ЛПР: не хватает мощности, на основе прогнозных значений потребления электроэнергии автоматически подключаются линии перетока электроэнергии из соседних участков (районов обслуживания), а избыток мощности перенаправляется в необходимые участки (районы обслуживания).

2. Δ_2 – ситуация, требующая вмешательства ЛПР, но не требующая подключения эксперта #1: случай, когда требуются необходимые правила фильтрации или прогнозные значения для принятия управленческих решений. Например, отклонения, требующие вмешательства ЛПР: появилась необходимость

в организации дополнительной выработки электроэнергии на основе прогнозных значений потребления электроэнергии.

3. Δ_3 – ситуация, требующая подключения эксперта #1: случай, когда требуются прогнозные значения потребления электроэнергии для принятия управленческих решений, при этом параметры ИНС настроены некорректно, как следствие, меньшая прогнозная точность значений на выходе. Например, корректируем весовые коэффициенты, благодаря чему повышаем прогнозную точность, как следствие, улучшаем качество управления электросетями и нагрузкой на них.

4. Δ_4 – ситуация, требующая подключение эксперта #2 в двух случаях: при существенном изменении сетевого потока данных (внешнее возмущающее воздействие); при некорректном формировании правила обработки информации (сетевой поток данных). Например, руководство предприятия обновило датчики и счетчик, в которых данные передаются по новому протоколу и с разным типом и видом сетевого потока данных.

Таблица 4.10 – Фрагмент исходного датасета, сформированного импульсной нейронной сетью

Дата ($D^k(m)$)	Номер счет- чика (k)	Потреб- ляемая электро- энергия, кВт*ч (EE_{ji}^k)	Темпе- ратура воздуха, °С (AT_{ji}^k)	Количест- во осадков, мм (Pr_{ji}^k)	Выходной день	Сезон ($Season_{ji}^k$)
01.01.2018	1	7156,5	-12	0	1	1
02.01.2018		7567,1	-14		1	1
03.01.2018		8123,2	-15		1	1
04.01.2018		8119,3	-15		1	1
05.01.2018		8912,0	-16		1	1
06.01.2018		7118,4	-15		1	1
07.01.2018		7945,2	-16		1	1
08.01.2018		7981,5	-16		1	1
09.01.2018		8589,0	-15		1	1
10.01.2018		8611,7	-17		0	1

В ходе эксперимента проведено большое количество расчетов и итераций в рамках прогнозирования количества потребляемой электроэнергии. Общий массив используемых в эксперименте данных составляет 25 903 значения. Из этих 25 903 значений 60% подаются на обучающую выборку нейронной сети, на тестовую и проверочную выборки – по 20%.

В данной диссертации приведены результаты эксперимента в рамках прогнозирования количества потребляемой электроэнергии клиентами (потребителями) в период с 01.04.2022 по 01.05.2022 (данные взяты с ограниченного числа счетчиков и датчиков). Изначально были рассчитаны прогнозные значения с их последующим сравнением с реальными (фактическими) значениями количества потребляемой электроэнергии (таблицы 4.11–4.12, рисунок 4.2).

Ежедневно (каждые сутки в 8:00 утра) данные снимаются со всех электросчетчиков (на текущий момент их количество составляет 196, т.е. $n = 196$), и заносятся в базу данных для последующей обработки с целью расчета прогнозных значений с помощью искусственной нейронной сети. Датой начала эксперимента можно считать 01.04.2022, т.е. $i=1$.

Сам прогноз осуществлялся на следующие m (например, 1–31) дней по 1 электросчетчику, т.е. $n=1$. Такое значение n взято для сокращения времени эксперимента. На следующий день измерялось фактическое значение потребления количества электроэнергии на этих же электросчетчиках $EE_{j,i+m}^k$. ЛПП (рисунок 2.2), например, начальник отдела, в зависимости от требований предприятия, выбирает количество дней, на которые будет осуществляться прогноз. Осуществляя прогнозирование ежедневно до 01.05.2022 включительно, получаем по каждому электросчетчику 961 прогноз и фактическое значение количества потребляемой электроэнергии по каждому из n электросчетчиков. Для этого введем дополнительное обозначение l – количество дней эксперимента, то есть значение $l_{k,i+m} = 31$. Это на основании формулы (2.27) дает возможность определить

среднюю точность (определить среднеквадратичную ошибку E_{cp}^k) прогноза по каждому k -му электросчетчику:

$$E_{cp}^k = \sqrt{\frac{1}{l_{k,i+m}} \sum_{l=1}^{l_{k,i+m}} |(\overline{Ee_{jl}^k} - \overline{Ee_{jl}^k})|^2}, \quad (4.8)$$

где $\overline{Ee_{ji}^k}$ желаемое (эталонное) значение выходных векторов, а выходы сети представлены в качестве вектора выхода ($\overline{EE_{ji}^k}$). При этом ошибка в процентах может быть определена по соотношению (4.9), где в знаменателе стоит среднее значение векторов выхода за время эксперимента:

$$Ep^k = \left(\frac{E_{cp}^k}{\frac{1}{l_{k,i+m}} \sum_{l=1}^{l_{k,i+m}} \overline{Ee_{jl}^k}} \right) \cdot 100\%. \quad (4.9)$$

Таблица 4.11 – Прогнозные значения суточного количества потребляемой энергии клиентами (на примере предприятия «Кумертауские электрические сети», счетчик №1), полученные на основе расчета искусственной нейронной сетью на 1–24 дня

#	Дата прогноза	Прогноз на 4 дня, кВт*ч	Прогноз на 3 дня, кВт*ч	Прогноз на 2 дня, кВт*ч	Прогноз на 1 день, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	6421,0	6429,1	6420,2	6419,0	На 01.05.2022: 6422,8
2	02.04.2022	6178,1	6259,3	6203,4	6173,9	На 02.05.2022: 6189,2
3	03.04.2022	6601,2	6651,3	6649,0	6622,1	На 03.05.2022: 6625,2
4	04.04.2022	6807,2	6716,6	6750,5	6751,8	На 04.05.2022: 6803,0
5	05.04.2022	7480,5	7501,0	7510,3	7519,8	На 05.05.2022: 7519,8
6	06.04.2022	7755,8	7720,4	7729,9	7764,3	На 06.05.2022: 7785,4
7	07.04.2022	6715,4	6845,3	6866,6	6910,8	На 07.05.2022: 6962,4
8	08.04.2022	6939,0	6949,1	6982,3	7003,0	На 08.05.2022: 7005,8
9	09.04.2022	6631,0	6639,4	6645,0	6655,1	На 09.05.2022: 6664,2
10	10.04.2022	6612,3	6678,0	6680,3	6690,7	На 10.05.2022: 6665,0
11	11.04.2022	7090,1	6965,4	6990,4	7153,8	На 11.05.2022: 7175,8

Таблица 4.11 – Продолжение

12	12.04.2022	6618,0	6617,2	6623,1	6630.1	На 12.05.2022: 6630,4
13	13.04.2022	6920,2	6918,3	6910,3	6903.4	На 13.05.2022: 6883,8
14	14.04.2022	6570,3	6560,1	6572,4	6575.3	На 14.05.2022: 6576,0
15	15.04.2022	6618,9	6634,2	6655,5	6670.0	На 15.05.2022: 6673,8
16	16.04.2022	6634,5	6641,0	6655,3	6679.0	На 16.05.2022: 6689,6
17	17.04.2022	7180,0	7183,2	7160,3	7201.3	На 17.05.2022: 7208,2
18	18.04.2022	7221,5	7255,8	7290,1	7307.8	На 18.05.2022: 7310,8
19	19.04.2022	7220,0	7161,0	7203,0	7223.0	На 19.05.2022: 7223,8
20	20.04.2022	7206,7	7228,0	7287,1	7333.8	На 20.05.2022: 7333,8
21	21.04.2022	6627,3	6619,1	6688,2	6639.1	На 21.05.2022: 6630,4
22	22.04.2022	6490,4	6500,1	6580,3	6580.4	На 22.05.2022: 6679,4
23	23.04.2022	7438,0	7451,7	7349,3	7560.3	На 23.05.2022: 7491,2
24	24.04.2022	6840,0	6891,7	6854,3	6855.2	На 24.05.2022: 6977,4
25	25.04.2022	6553,0	6619,8	6570,1	6584.2	На 25.05.2022: 6594,4
26	26.04.2022	6598,8	6651,7	6704,3	6755.7	На 26.05.2022: 6828,8
27	27.04.2022	6835,3	6906,0	7149,1	7273.7	На 27.05.2022: 6939,6
28	28.04.2022	6490,1	6498,3	6513,5	6529.4	На 28.05.2022: 6529,8
29	29.04.2022	6013,8	6113,0	6120,1	6098.5	На 29.05.2022: 6019,8
30	30.04.2022	5900,2	5912,2	5894,1	6019.0	На 30.05.2022: 6027,0
31	01.05.2022	5800,0	5812,5	5771,4	5800.1	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		62,7	61,4	57,2	34,2	
Средняя ошибка, %		1,08	1,06	0,99	0,59	
#	Дата прогноза	Прогноз на 8 дней, кВт*ч	Прогноз на 7 дней, кВт*ч	Прогноз на 6 дней, кВт*ч	Прогноз на 5 дней, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	6281,5	6380,9	6395,0	6410,0	На 01.05.2022: 6422,8
2	02.04.2022	6259,6	6267,0	6290,3	6217,2	На 02.05.2022: 6189,2
3	03.04.2022	6306,1	6478,7	6479,2	6589,0	На 03.05.2022: 6625,2
4	04.04.2022	6345,2	6731,5	6349,8	6757,3	На 04.05.2022: 6803,0
5	05.04.2022	7297,3	7411,5	7425,2	7435,1	На 05.05.2022: 7519,8
6	06.04.2022	7680,0	7538,9	7708,5	7756,3	На 06.05.2022: 7785,4
7	07.04.2022	6600,7	6645,9	6682,4	6702,8	На 07.05.2022: 6962,4
8	08.04.2022	6894,1	6912,3	6926,5	6935,3	На 08.05.2022: 7005,8
9	09.04.2022	6480,7	6494,0	6538,0	6655,2	На 09.05.2022: 6664,2

Таблица 4.11 – Продолжение

10	10.04.2022	6526,2	6571,4	6580,1	6592,6	На 10.05.2022: 6665,0
11	11.04.2022	7046,0	7097,0	7069,1	6917,1	На 11.05.2022: 7175,8
12	12.04.2022	6513,2	6597,1	6603,3	6612,5	На 12.05.2022: 6630,4
13	13.04.2022	6738,1	6800,2	6811,3	6843,4	На 13.05.2022: 6883,8
14	14.04.2022	6482,1	6551,1	6560,1	6589,2	На 14.05.2022: 6576,0
15	15.04.2022	6486,0	6509,2	6577,1	6612,3	На 15.05.2022: 6673,8
16	16.04.2022	6578,0	6580,0	6121,3	6632,2	На 16.05.2022: 6689,6
17	17.04.2022	7006,3	7100,2	7145,3	7162,4	На 17.05.2022: 7208,2
18	18.04.2022	7183,2	7189,0	7201,4	7203,2	На 18.05.2022: 7310,8
19	19.04.2022	7169,0	7176,0	7152,0	7212,0	На 19.05.2022: 7223,8
20	20.04.2022	7202,1	7220,2	7217,1	7286,6	На 20.05.2022: 7333,8
21	21.04.2022	6653,2	6656,1	6644,6	6667,0	На 21.05.2022: 6630,4
22	22.04.2022	6658,2	6210,4	6545,7	6488,1	На 22.05.2022: 6679,4
23	23.04.2022	7332,2	7475,3	7554,0	7481,2	На 23.05.2022: 7491,2
24	24.04.2022	6822,8	6835,1	6810,0	6851,2	На 24.05.2022: 6977,4
25	25.04.2022	6534,6	6482,2	6550,0	6483,3	На 25.05.2022: 6594,4
26	26.04.2022	6541,6	6598,3	6633,1	6560,2	На 26.05.2022: 6828,8
27	27.04.2022	6853,1	6760,4	6843,1	6870,5	На 27.05.2022: 6939,6
28	28.04.2022	6470,2	6473,4	6443,2	6463,9	На 28.05.2022: 6529,8
29	29.04.2022	6018,4	6014,2	5986,8	6012,7	На 29.05.2022: 6019,8
30	30.04.2022	5861,0	5905,7	5898,2	5924,1	На 30.05.2022: 6027,0
31	01.05.2022	5993,0	5897,2	5428,1	5848,0	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		149,4	119,8	130,5	75,9	
Средняя ошибка, %		2,58	2,06	2,25	1,31	
#	Дата прогноза	Прогноз на 12 дней, кВт*ч	Прогноз на 11 дней, кВт*ч	Прогноз на 10 дней, кВт*ч	Прогноз на 9 дней, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	5972,3	6005,3	6159,0	6205,2	На 01.05.2022: 6422,8
2	02.04.2022	5888,0	6100,5	6112,9	6297,3	На 02.05.2022: 6189,2
3	03.04.2022	5959,0	5901,7	6176,2	6294,2	На 03.05.2022: 6625,2
4	04.04.2022	5997,3	5999,1	6411,3	6395,0	На 04.05.2022: 6803,0
5	05.04.2022	7129,2	7200,0	7389,3	7400,0	На 05.05.2022: 7519,8
6	06.04.2022	7175,1	7350,0	7421,5	7586,3	На 06.05.2022: 7785,4
7	07.04.2022	6199,1	6300,3	6412,0	6520,2	На 07.05.2022: 6962,4

Таблица 4.11 – Продолжение

8	08.04.2022	6786,3	6900,9	6851,1	6877,7	На 08.05.2022: 7005,8
9	09.04.2022	6090,4	6279,0	6300,0	6442,2	На 09.05.2022: 6664,2
10	10.04.2022	6272,4	6376,0	6412,9	6482,0	На 10.05.2022: 6665,0
11	11.04.2022	6791,5	6857,5	6912,8	6947,3	На 11.05.2022: 7175,8
12	12.04.2022	6099,6	6248,3	6387,9	6521,3	На 12.05.2022: 6630,4
13	13.04.2022	6186,7	6321,4	6389,5	6668,2	На 13.05.2022: 6883,8
14	14.04.2022	6132,3	6314,5	6398,7	6455,2	На 14.05.2022: 6576,0
15	15.04.2022	6141,6	6276,3	6420,7	6497,3	На 15.05.2022: 6673,8
16	16.04.2022	6248,7	6164,0	6450,6	6559,2	На 16.05.2022: 6689,6
17	17.04.2022	6717,9	6811,9	6850,3	6940,0	На 17.05.2022: 7208,2
18	18.04.2022	7038,1	7120,0	7150,6	7160,3	На 18.05.2022: 7310,8
19	19.04.2022	7049,6	7230,1	7089,2	7221,0	На 19.05.2022: 7223,8
20	20.04.2022	6945,3	7000,2	7143,5	7278,1	На 20.05.2022: 7333,8
21	21.04.2022	6596,8	6800,1	6715,4	6654,8	На 21.05.2022: 6630,4
22	22.04.2022	5998,4	6047,0	6120,5	6573,3	На 22.05.2022: 6679,4
23	23.04.2022	7070,8	7210,1	7250,4	7435,4	На 23.05.2022: 7491,2
24	24.04.2022	6955,1	7101,0	6789,1	6842,0	На 24.05.2022: 6977,4
25	25.04.2022	6110,2	6222,0	6310,8	6532,0	На 25.05.2022: 6594,4
26	26.04.2022	6135,6	6259,0	6270,3	6480,4	На 26.05.2022: 6828,8
27	27.04.2022	6009,5	6353,0	6413,7	6756,0	На 27.05.2022: 6939,6
28	28.04.2022	6000,9	6242,0	6380,1	6484,1	На 28.05.2022: 6529,8
29	29.04.2022	5941,3	5952,0	5960,7	5996,1	На 29.05.2022: 6019,8
30	30.04.2022	5997,1	6084,0	5812,4	5828,8	На 30.05.2022: 6027,0
31	01.05.2022	6120,1	6407,0	6299,8	6024,3	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		443,8	366,6	286,4	168,7	
Средняя ошибка, %		7,65	6,32	4,94	2,91	
#	Дата прогноза	Прогноз на 16 дней, кВт*ч	Прогноз на 15 дней, кВт*ч	Прогноз на 14 дней, кВт*ч	Прогноз на 13 дней, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	5791,3	5832,0	5900,0	5970,1	На 01.05.2022: 6422,8
2	02.04.2022	5813,6	5768,0	5885,0	5681,0	На 02.05.2022: 6189,2
3	03.04.2022	5910,1	5943,0	5970,0	5949,5	На 03.05.2022: 6625,2
4	04.04.2022	5897,3	5904,1	5932,2	5947,1	На 04.05.2022: 6803,0
5	05.04.2022	7021,4	7030,4	7073,1	7078,1	На 05.05.2022: 7519,8

Таблица 4.11 – Продолжение

6	06.04.2022	7131,5	7137,2	7145,1	7168,2	На 06.05.2022: 7785,4
7	07.04.2022	6102,0	6172,3	6191,3	6194,5	На 07.05.2022: 6962,4
8	08.04.2022	6676,2	6698,4	6715,5	6764,7	На 08.05.2022: 7005,8
9	09.04.2022	6024,2	6041,5	6045,7	6054,1	На 09.05.2022: 6664,2
10	10.04.2022	6076,5	6149,5	6178,6	6220,2	На 10.05.2022: 6665,0
11	11.04.2022	6743,6	6753,6	6758,6	6763,4	На 11.05.2022: 7175,8
12	12.04.2022	6015,2	6022,1	6026,1	6044,3	На 12.05.2022: 6630,4
13	13.04.2022	6165,7	6175,2	6182,1	6183,8	На 13.05.2022: 6883,8
14	14.04.2022	6066,8	6068,4	6069,0	6120,8	На 14.05.2022: 6576,0
15	15.04.2022	6123,3	6124,2	6136,8	6138,7	На 15.05.2022: 6673,8
16	16.04.2022	6179,3	6188,3	6203,1	6218,5	На 16.05.2022: 6689,6
17	17.04.2022	6687,4	6693,8	6697,2	6712,3	На 17.05.2022: 7208,2
18	18.04.2022	7012,6	7019,8	7026,3	7028,4	На 18.05.2022: 7310,8
19	19.04.2022	7034,7	7038,9	7039,5	7042,5	На 19.05.2022: 7223,8
20	20.04.2022	6900,8	6923,1	6927,5	6934,3	На 20.05.2022: 7333,8
21	21.04.2022	6570,1	6574,1	6588,3	6592,5	На 21.05.2022: 6630,4
22	22.04.2022	5944,1	5973,0	5979,4	5984,6	На 22.05.2022: 6679,4
23	23.04.2022	7010,4	7016,5	7042,6	7065,7	На 23.05.2022: 7491,2
24	24.04.2022	6888,6	6894,6	6934,7	6936,4	На 24.05.2022: 6977,4
25	25.04.2022	6032,7	6035,5	6051,7	6068,1	На 25.05.2022: 6594,4
26	26.04.2022	6053,9	6057,3	6077,8	6094,1	На 26.05.2022: 6828,8
27	27.04.2022	5993,1	5997,3	6000,1	6004,6	На 27.05.2022: 6939,6
28	28.04.2022	5891,2	5957,4	5958,2	5970,6	На 28.05.2022: 6529,8
29	29.04.2022	5886,3	5896,5	5923,3	5937,3	На 29.05.2022: 6019,8
30	30.04.2022	5878,5	5970,9	5977,4	5983,5	На 30.05.2022: 6027,0
31	01.05.2022	6101,0	6111,0	6114,4	6115,5	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		511,3	494,4	475,8	468,9	
Средняя ошибка, %		8,81	8,52	8,20	8,08	
#	Дата прогноза	Прогноз на 20 дней, кВт*ч	Прогноз на 19 дней, кВт*ч	Прогноз на 18 дней, кВт*ч	Прогноз на 17 дней, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	5812,6	5914,2	5943,0	5600,0	На 01.05.2022: 6422,8
2	02.04.2022	5814,4	5837,0	5890,0	5915,1	На 02.05.2022: 6189,2
3	03.04.2022	5830,3	5850,1	5901,0	5909,0	На 03.05.2022: 6625,2

Таблица 4.11 – Продолжение

4	04.04.2022	5840,3	5858,3	5880,1	5890,2	На 04.05.2022: 6803,0
5	05.04.2022	6976,6	6981,3	6987,3	6995,4	На 05.05.2022: 7519,8
6	06.04.2022	7008,9	7026,9	7048,5	7115,4	На 06.05.2022: 7785,4
7	07.04.2022	6030,7	6052,5	6084,5	6091,5	На 07.05.2022: 6962,4
8	08.04.2022	6623,7	6639,3	6648,6	6664,6	На 08.05.2022: 7005,8
9	09.04.2022	5964,2	5984,5	6002,2	6005,1	На 09.05.2022: 6664,2
10	10.04.2022	5945,7	5950,5	5977,3	6042,1	На 10.05.2022: 6665,0
11	11.04.2022	6668,3	6679,0	6690,1	6708,3	На 11.05.2022: 7175,8
12	12.04.2022	5906,6	5977,2	5997,6	6007,1	На 12.05.2022: 6630,4
13	13.04.2022	6115,2	6128,3	6129,7	6139,4	На 13.05.2022: 6883,8
14	14.04.2022	6022,2	6033,4	6036,8	6065,5	На 14.05.2022: 6576,0
15	15.04.2022	6094,3	6097,5	6110,1	6122,6	На 15.05.2022: 6673,8
16	16.04.2022	6121,6	6144,7	6159,2	6164,7	На 16.05.2022: 6689,6
17	17.04.2022	6650,8	6660,7	6661,3	6666,8	На 17.05.2022: 7208,2
18	18.04.2022	6939,1	6962,9	6983,5	6984,1	На 18.05.2022: 7310,8
19	19.04.2022	6916,1	6957,0	6964,5	6990,1	На 19.05.2022: 7223,8
20	20.04.2022	6870,7	6875,3	6879,6	6881,2	На 20.05.2022: 7333,8
21	21.04.2022	6545,0	6556,3	6559,7	6568,3	На 21.05.2022: 6630,4
22	22.04.2022	5863,0	5876,1	5911,2	5935,4	На 22.05.2022: 6679,4
23	23.04.2022	6904,0	6937,2	6955,2	7002,5	На 23.05.2022: 7491,2
24	24.04.2022	6822,7	6840,3	6883,3	6885,4	На 24.05.2022: 6977,4
25	25.04.2022	5971,4	5977,7	5990,4	6025,4	На 25.05.2022: 6594,4
26	26.04.2022	6022,5	6027,8	6033,4	6044,4	На 26.05.2022: 6828,8
27	27.04.2022	5943,7	5945,5	5952,5	5985,5	На 27.05.2022: 6939,6
28	28.04.2022	5842,2	5852,2	5854,8	5884,6	На 28.05.2022: 6529,8
29	29.04.2022	5832,4	5845,3	5855,1	5881,7	На 29.05.2022: 6019,8
30	30.04.2022	5795,7	5827,4	5848,1	5861,7	На 30.05.2022: 6027,0
31	01.05.2022	6033,1	6042,5	6050,1	6067,8	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		568,1	549,0	532,4	526,0	
Средняя ошибка, %		9,79	9,46	9,18	9,07	
#	Дата прогноза	Прогноз на 24 дня, кВт*ч	Прогноз на 23 дня, кВт*ч	Прогноз на 22 дня, кВт*ч	Прогноз на 21 день, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	5631,1	5644,1	5772,3	5803,8	На 01.05.2022: 6422,8

Таблица 4.11 – Окончание

2	02.04.2022	5603,3	5640,4	5654,2	5788,3	На 02.05.2022: 6189,2
3	03.04.2022	5626,5	5650,0	5763,1	5812,4	На 03.05.2022: 6625,2
4	04.04.2022	5724,6	5744,0	5819,2	5831,7	На 04.05.2022: 6803,0
5	05.04.2022	6780,5	6815,3	6935,3	6963,5	На 05.05.2022: 7519,8
6	06.04.2022	6944,4	6947,3	7001,5	7005,5	На 06.05.2022: 7785,4
7	07.04.2022	5980,3	5987,6	6004,4	6005,7	На 07.05.2022: 6962,4
8	08.04.2022	6370,2	6422,5	6605,6	6612,5	На 08.05.2022: 7005,8
9	09.04.2022	5880,8	5886,4	5948,6	5953,5	На 09.05.2022: 6664,2
10	10.04.2022	5839,8	5860,4	5921,8	5927,5	На 10.05.2022: 6665,0
11	11.04.2022	6546,0	6556,3	6647,0	6652,7	На 11.05.2022: 7175,8
12	12.04.2022	5777,9	5795,1	5893,0	5900,4	На 12.05.2022: 6630,4
13	13.04.2022	5963,9	6000,2	6106,2	6111,4	На 13.05.2022: 6883,8
14	14.04.2022	5883,5	5885,5	6008,1	6012,7	На 14.05.2022: 6576,0
15	15.04.2022	5922,3	5939,4	6030,2	6044,5	На 15.05.2022: 6673,8
16	16.04.2022	5831,3	5928,5	6050,0	6113,1	На 16.05.2022: 6689,6
17	17.04.2022	6566,5	6570,9	6602,5	6619,7	На 17.05.2022: 7208,2
18	18.04.2022	6785,6	6799,0	6922,0	6924,0	На 18.05.2022: 7310,8
19	19.04.2022	6786,8	6793,8	6904,0	6915,2	На 19.05.2022: 7223,8
20	20.04.2022	6744,1	6764,7	6807,1	6843,0	На 20.05.2022: 7333,8
21	21.04.2022	6405,1	6470,6	6539,3	6544,9	На 21.05.2022: 6630,4
22	22.04.2022	5692,3	5743,5	5805,8	5818,6	На 22.05.2022: 6679,4
23	23.04.2022	6711,6	6754,3	6851,7	6862,3	На 23.05.2022: 7491,2
24	24.04.2022	6647,2	6689,2	6797,0	6800,7	На 24.05.2022: 6977,4
25	25.04.2022	5850,2	5881,1	5938,0	5957,5	На 25.05.2022: 6594,4
26	26.04.2022	5847,3	5921,1	6016,0	6021,1	На 26.05.2022: 6828,8
27	27.04.2022	5766,4	5771,3	5901,5	5941,1	На 27.05.2022: 6939,6
28	28.04.2022	5777,5	5794,4	5812,3	5813,7	На 28.05.2022: 6529,8
29	29.04.2022	5695,8	5750,5	5826,2	5830,1	На 29.05.2022: 6019,8
30	30.04.2022	5633,5	5658,5	5784,1	5791,7	На 30.05.2022: 6027,0
31	01.05.2022	5877,7	5897,6	6003,0	6029,0	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		707,5	680,8	600,2	583,2	
Средняя ошибка, %		12,19	11,73	10,34	10,05	

Таблица 4.12 – Прогнозные значения суточного количества потребляемой энергии клиентами (на примере предприятия «Кумертауские электрические сети», счетчик №1), полученные на основе расчета искусственной нейронной сетью на 25–31 дни

#	Дата прогноза	Прогноз на 28 дней, кВт*ч	Прогноз на 27 дней, кВт*ч	Прогноз на 26 дней, кВт*ч	Прогноз на 25 дней, кВт*ч	Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	5571,0	5591,3	5610,1	5620,3	На 01.05.2022: 6422,8
2	02.04.2022	5488,0	5510,5	5529,2	5585,4	На 02.05.2022: 6189,2
3	03.04.2022	5549,2	5555,4	5585,4	5623,6	На 03.05.2022: 6625,2
4	04.04.2022	5671,4	5691,3	5693,5	5713,8	На 04.05.2022: 6803,0
5	05.04.2022	6743,6	6748,2	6750,6	6759,1	На 05.05.2022: 7519,8
6	06.04.2022	6860,3	6925,4	6926,6	6935,7	На 06.05.2022: 7785,4
7	07.04.2022	5937,1	5952,6	5967,3	5971,3	На 07.05.2022: 6962,4
8	08.04.2022	6351,4	6355,7	6362,4	6365,0	На 08.05.2022: 7005,8
9	09.04.2022	5790,6	5849,9	5855,7	5856,3	На 09.05.2022: 6664,2
10	10.04.2022	5787,8	5811,1	5816,6	5838,4	На 10.05.2022: 6665,0
11	11.04.2022	6426,5	6455,2	6509,5	6539,1	На 11.05.2022: 7175,8
12	12.04.2022	5745,4	5755,4	5762,9	5776,3	На 12.05.2022: 6630,4
13	13.04.2022	5906,3	5929,6	5939,0	5952,3	На 13.05.2022: 6883,8
14	14.04.2022	5803,2	5807,7	5812,1	5815,4	На 14.05.2022: 6576,0
15	15.04.2022	5786,5	5814,5	5817,2	5920,6	На 15.05.2022: 6673,8
16	16.04.2022	5734,6	5744,3	5758,3	5794,5	На 16.05.2022: 6689,6
17	17.04.2022	6473,8	6527,5	6534,5	6558,3	На 17.05.2022: 7208,2
18	18.04.2022	6722,9	6752,3	6780,7	6783,1	На 18.05.2022: 7310,8
19	19.04.2022	6751,4	6756,1	6758,8	6774,0	На 19.05.2022: 7223,8
20	20.04.2022	6669,3	6688,4	6712,2	6736,9	На 20.05.2022: 7333,8
21	21.04.2022	6306,3	6368,6	6387,3	6395,8	На 21.05.2022: 6630,4
22	22.04.2022	5537,1	5609,7	5661,6	5686,1	На 22.05.2022: 6679,4
23	23.04.2022	6616,2	6653,4	6663,8	6707,3	На 23.05.2022: 7491,2
24	24.04.2022	6513,5	6555,3	6596,6	6635,4	На 24.05.2022: 6977,4
25	25.04.2022	5746,4	5764,8	5777,7	5784,5	На 25.05.2022: 6594,4
26	26.04.2022	5781,8	5791,4	5792,1	5834,3	На 26.05.2022: 6828,8
27	27.04.2022	5628,3	5713,3	5733,5	5746,2	На 27.05.2022: 6939,6
28	28.04.2022	5702,4	5708,1	5722,4	5751,7	На 28.05.2022: 6529,8
29	29.04.2022	5524,1	5553,8	5578,3	5652,5	На 29.05.2022: 6019,8
30	30.04.2022	5545,3	5563,4	5587,3	5593,4	На 30.05.2022: 6027,0

Таблица 4.12 – Продолжение

31	01.05.2022	5795,6	5815,3	5837,1	5872,1	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		787,0	760,8	746,0	723,8	
Средняя ошибка, %		13,56	13,11	12,86	12,48	
#	Дата прогноза	Прогноз на 31 день, кВт*ч	Прогноз на 30 дней, кВт*ч	Прогноз на 29 дней, кВт*ч		Фактическое значение количества потребляемой энергии на прогнозируемую дату, кВт*ч
1	01.04.2022	5519,0	5532,1	5552,1		На 01.05.2022: 6422,8
2	02.04.2022	5403,1	5418,2	5435,2		На 02.05.2022: 6189,2
3	03.04.2022	5468,1	5500,5	5544,4		На 03.05.2022: 6625,2
4	04.04.2022	5635,5	5638,6	5652,5		На 04.05.2022: 6803,0
5	05.04.2022	6671,7	6693,7	6707,4		На 05.05.2022: 7519,8
6	06.04.2022	6809,6	6821,8	6837,3		На 06.05.2022: 7785,4
7	07.04.2022	5862,3	5881,3	5931,6		На 07.05.2022: 6962,4
8	08.04.2022	6301,4	6316,3	6332,7		На 08.05.2022: 7005,8
9	09.04.2022	5756,5	5774,3	5785,8		На 09.05.2022: 6664,2
10	10.04.2022	5701,9	5738,4	5779,1		На 10.05.2022: 6665,0
11	11.04.2022	6312,5	6319,5	6411,2		На 11.05.2022: 7175,8
12	12.04.2022	5664,1	5713,6	5714,5		На 12.05.2022: 6630,4
13	13.04.2022	5823,5	5869,4	5877,7		На 13.05.2022: 6883,8
14	14.04.2022	5732,3	5734,2	5761,8		На 14.05.2022: 6576,0
15	15.04.2022	5620,4	5629,6	5773,9		На 15.05.2022: 6673,8
16	16.04.2022	5628,5	5636,1	5651,0		На 16.05.2022: 6689,6
17	17.04.2022	6356,2	6370,6	6467,0		На 17.05.2022: 7208,2
18	18.04.2022	6650,2	6670,7	6682,1		На 18.05.2022: 7310,8
19	19.04.2022	6700,1	6723,7	6732,2		На 19.05.2022: 7223,8
20	20.04.2022	6514,0	6591,3	6609,3		На 20.05.2022: 7333,8
21	21.04.2022	6226,8	6282,2	6297,3		На 21.05.2022: 6630,4
22	22.04.2022	5504,1	5525,2	5530,5		На 22.05.2022: 6679,4
23	23.04.2022	6508,3	6516,4	6557,7		На 23.05.2022: 7491,2
24	24.04.2022	6479,5	6483,0	6504,9		На 24.05.2022: 6977,4
25	25.04.2022	5702,3	5714,6	5731,1		На 25.05.2022: 6594,4
26	26.04.2022	5666,5	5699,5	5743,4		На 26.05.2022: 6828,8
27	27.04.2022	5602,3	5608,4	5619,3		На 27.05.2022: 6939,6
28	28.04.2022	5612,2	5657,3	5667,3		На 28.05.2022: 6529,8

29	29.04.2022	5508,4	5514,3	5520,2	На 29.05.2022: 6019,8
30	30.04.2022	5431,5	5468,1	5512,5	На 30.05.2022: 6027,0
31	01.05.2022	5765,7	5770,1	5779,7	На 31.05.2022: 5792,6
Средняя ошибка, кВт*ч		861,9	840,2	811,2	
Средняя ошибка, %		14,86	14,48	13,98	

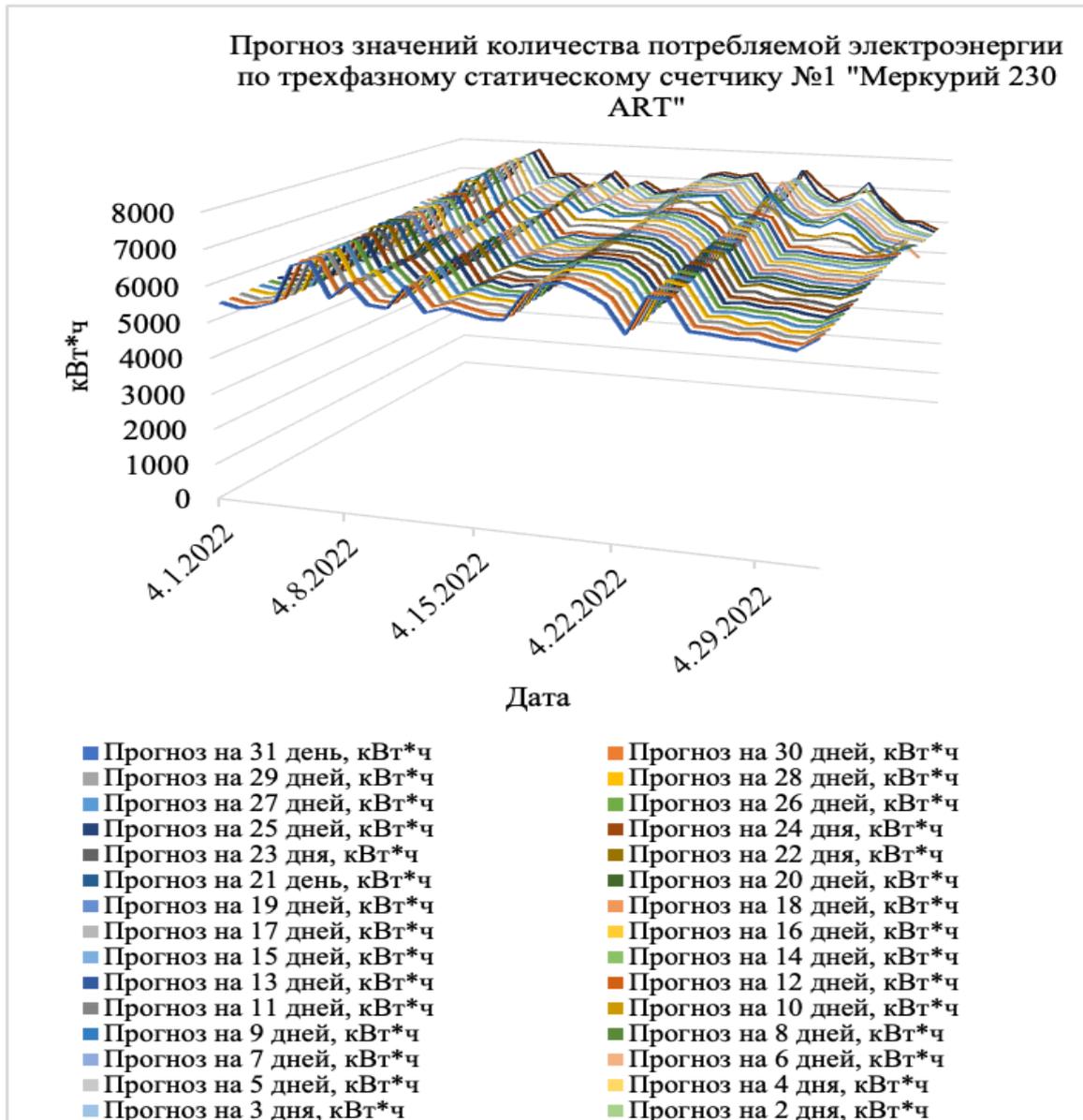


Рисунок 4.2 – Результаты прогнозирования суточного количества потребляемой электроэнергии в рамках проведенного эксперимента

Исходя из результатов, представленных в таблицах 4.11–4.12 и на рисунке 4.2, проведенного эксперимента можем сделать следующие выводы:

1) во-первых, средняя ошибка (рисунок 4.3 (а, б)) прогнозируемого значения потребления электроэнергии составила 457.2 кВт*ч (E_p^k – 7.8%) по всему трехфазному статическому счетчику №1 «Меркурий 230 ART»;

2) во-вторых, можно отметить, что в прогнозе на 6-ой день (таблица 4.11) ошибки E_{cp}^k и E_p^k больше, чем при прогнозе на 7-ой день. Это является ограничением со стороны модифицированного метода и модели для обучения искусственной нейронной сети. Но в связи с тем, что весь прототип СППР [177] (Приложение А) реализован на языках программирования Java и Python «с нуля», появляется «окно» для дальнейших исследований в рассматриваемой предметной области в целом, и по снятию данного ограничения в частности;

3) в-третьих, несмотря на слабые тенденции к возрастанию/понижению (рисунок 4.4) потребления электроэнергии, а также постоянные всплески (выбросы, аномалии) во временном ряде, прогнозные значения являются достаточно точными за счет введенных ранее коэффициентов (глава 2), что подтверждают дальнейшие эксперименты.

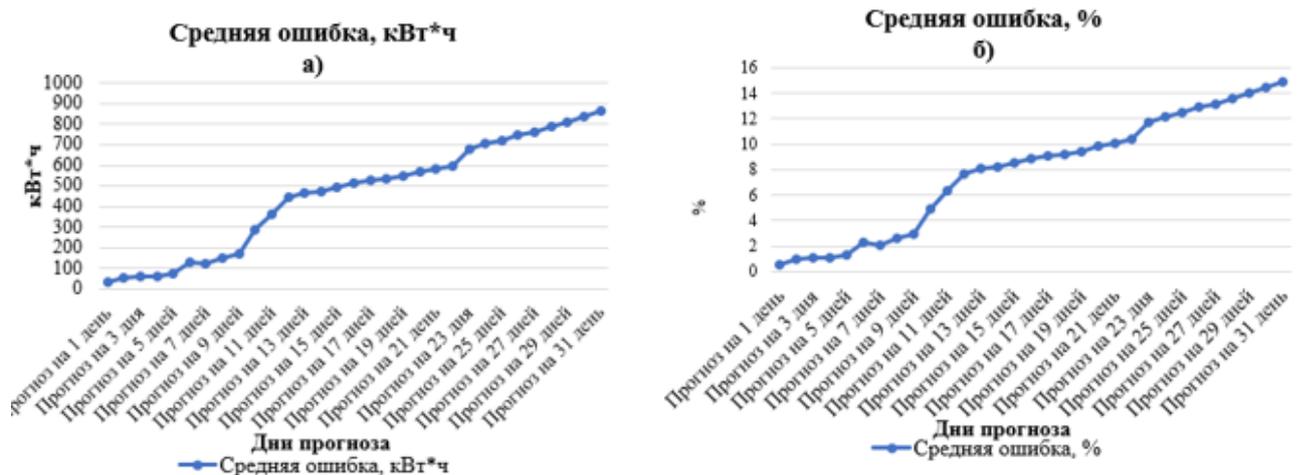


Рисунок 4.3 – а) Средние ошибки прогнозирования количества потребляемой электроэнергии

Рисунок 4.3 – б) Средние ошибки прогнозирования количества потребляемой электроэнергии



Рисунок 4.4 – Ретроспективные данные потребления электроэнергии за последние 4 года, используемые для проведения эксперимента

Также с целью дополнительного подтверждения эффективности предлагаемого решения, а также для дальнейших исследований в электроэнергетической области были проведены следующие эксперименты:

а) сравнительный анализ методов обучения. Суть эксперимента заключается в сравнении средних ошибок и скорости обучения при прогнозировании количества потребляемой электроэнергии ИНС классическим методом обратного распространения ошибки, предложенным в диссертации модифицированным методом, а также другими методами обучения;

б) сравнительный анализ предлагаемой РНС и статистических/гибридных методов прогнозирования. Суть эксперимента заключается в сравнении средних ошибок предлагаемой в диссертации ИНС со статистическими/гибридными методами прогнозирования, приведенными в главе 1;

с) сравнительный анализ предлагаемой РНС и аналогичных решений. Суть эксперимента заключается в сравнении средних ошибок и времени обучения предлагаемой в диссертации ИНС с другими искусственными нейронными сетями.

Все эксперименты проводились аналогично эксперименту, результаты которого представлены в таблицах 4.11–4.12. При этом для удобства в таблицы вынесены итоговые результаты в виде средних значений.

Сравнительный анализ методов обучения. Результаты представлены в таблице 4.13 (прогноз производился также на 31 день). Ошибки рассчитывались по формуле (4.9).

Таблица 4.13 – Средние значения ошибок при расчете прогнозных значений различными методами обучения

Метод обучения	Общее среднее значение ошибок по всем дням, кВт*ч	Общее среднее значение ошибок по всем дням, %	Скорость обучения, минут	Точность прогнозов, %	Сокращение времени обучения РНС относительно других методов, %
РНС	457,20	7,80	0,50	92,20	X
Обратного распространения ошибки (классический)	1572,30	23,10	3,00	76,90	83,00
GA (Genetic algorithm)	1290,40	18,90	10,30	81,10	95,00
Conjugate gradient method	1351,80	19,80	8,40	80,20	94,00
Kohonen's learning method	1400,10	20,50	9,20	79,50	94,50
Levenberg-Marquardt method	1739,60	25,50	4,00	74,50	87,00

Проводя сравнительный анализ результатов, представленных в таблице 4.13, приходим к выводу, что модифицированный метод прогнозирования временных рядов предложенной ИНС показывает себя более точным при расчете прогнозных

значений и быстрее при обучении. Для чистоты эксперимента все параметры задавались идентично тем, которые описаны в главе 2.

Повышение точности и сокращение времени обучения относительно других методов обучения варьировалась в диапазонах 11,10–17,70% и 83,00–95,00% соответственно. Такая эффективность достигается за счет добавленных ранее констант: скорости обучения, что снижает время обучения ИНС; чувствительности, что повышает точность прогнозных значений; и реагирования, что позволяет реагировать на изменения видов и типов сетевого потока данных, как следствие, вырабатывать адаптивные под каждый случай правила фильтрации сетевого потока данных.

Сравнительный анализ предложенной ИНС и статистических/гибридных методов прогнозирования. Результаты представлены в таблице 4.14 (прогноз производился также на 31 день). Ошибки рассчитывались по формуле (4.9).

Таблица 4.14 – Средние значения ошибок при расчете прогнозных значений статистическими/гибридными методами обучения

Метод прогнозирования	Общее среднее значение ошибок по всем дням, кВт*ч	Общее среднее значение ошибок по всем дням, %	Точность прогнозов, %
Предложенная ИНС	457,2	7,8	92,2
ARIMA	2187,5	32,1	67,9
SVR	2113,6	31,0	69,0
ARIMAX-GARCH	1784,5	26,2	73,8
MLR	2341,9	34,4	65,6

Результаты (таблица 4.14) проведенного эксперимента: за счет ранее описанных коэффициентов наиболее точным методом (средняя прогнозная точность достигает 92,2%) прогнозирования оказалась разработанная в рамках диссертации РНС. Более того, статистические и гибридные методы оказались более

ресурсозатратными (в плане вычислительных мощностей) и ресурсоемкими (в плане времени расчетов прогнозных значений).

Для программной реализации данных методов прогнозирования использовались специализированные библиотеки на языке Python: ARIMA – «pmdarima»; SVR – Scikit-learn с модулем «SVR», заложенным внутри SVM; ARIMAX-GARCH – разработан собственный модуль в соответствии с математическими моделями работы [131]; MLR – Scikit-learn с модулем «linear_model».

Сравнительный анализ метода РНС и аналогичных решений. Результаты представлены в таблице 4.15 (прогноз производился также на 31 день). Ошибки рассчитывались по формуле (4.9).

Таблица 4.15 – Средние значения ошибок при расчете прогнозных значений искусственными нейронными сетями

Нейронная сеть	Комбинация функций активации	Общее среднее значение ошибок по всем дням, кВт*ч	Общее среднее значение ошибок по всем дням, %	Точность прогнозов, %
Однослойный персептрон	Сигмоидальная	2983,5	43,8	92,2
Разработанная РНС	Входной слой и промежуточные слои: сигмоидальная; выходной слой: линейная.	457,2	7,8	56,2
Многослойный персептрон		1571,0	23,1	76,9
Сверточная		742,7	10,9	89,1
Классическая рекуррентная		1043,2	15,3	84,7

Исходя из проведенного эксперимента, результаты которого представлены в таблице 4.15, можно также сделать вывод об эффективности разработанной в рамках диссертации РНС. Указанная точность достигнута за счет модифицированных нейрона и метода обучения путем добавления

специализированных коэффициентов. При этом повышение точности прогнозных значений при обучении относительно других ИНС варьировалась в диапазоне 11,10–17,70%. Важно отметить, что при любом другом наборе функций активации, независимо от их комбинации, прогнозные значения становятся менее точными.

Программная реализация всех нейронных сетей выполнялась на высокоуровневом языке программирования Python с использованием библиотеки TensorFlow (за исключением разработанной РНС – Java).

Все эксперименты проводились на следующем оборудовании:

1. Количество физических серверов — 30.
2. Процессор Intel Xeon 5690 (CPU — 60, физических ядер — 360, количество потоков — 720).
3. Оперативная память — 960 ГБ.
4. SSD — RAID 10 (NVME Samsung 980 pro 500ГБ каждый);
5. Внешний сетевой канал — 20 Гбит/с.
6. Внутренний сетевой канал (локальная сеть) — 100Гбит/с.

Исходя из вышеприведенных результатов экспериментов, разработанная в рамках диссертации искусственная нейронная сеть показала наиболее лучшие характеристики в плане точности и скорости обучения. Например, относительно других методов обучения (таблица 4.13) повышение точности прогноза и скорости обучения достигало 17,70% и 95,00% соответственно.

Важно отметить, что выбор модели и метода обучения искусственной нейронной сети, в отличие от статистических/гибридных решений, для прогнозирования потребления электроэнергии был сделан на основе широкой возможности их модификаций, что позволяет достичь более существенной точности выходных значений, а также повысить скорость обучения.

Таким образом, реализованная в рамках диссертации искусственная нейронная сеть показала более точный и стабильный результат, что доказывает целесообразность применения предлагаемого решения в рамках прогнозирования потребления электроэнергии. За счет данных прогнозов у предприятий появится

возможность проводить энергосберегающие мероприятия. Это позволит более эффективно использовать финансовые ресурсы энергетическим компаниям, что положительно скажется не только на тарифных планах электроэнергии для физических и юридических лиц, но и предоставит возможность оптимизации бюджета энергетической компании.

4.3. Оценка эффективности отдельных модулей прототипа СППР

В данном подразделе будет оценена эффективность разработанного прототипа СППР путем апробации системы поддержки принятия решений в целом и каждого из ее модулей в рамках нагрузки на ресурсы физического сервера.

Изначально планируется апробация удаленного клиента RCM (Remote Client Manager), отвечающего за первичную обработку входящего потока информации по всем физическим серверам кластера. Далее оценивается общая нагрузка модулями ИмНС, РНС и MySQL при обработке больших разнородных и темпоральных данных.

4.3.1. Удаленный клиент (RCM)

Нагрузка на физические ресурсы без активированных модулей (при включенном удаленном клиенте) представлена в таблице 4.16. В данном и последующих подпунктах информация о нагрузке берется из веб-интерфейса панели управления прототипом СППР и мониторинга Zabbix [180] (open-source software).

Таблица 4.16 – Нагрузка на физические ресурсы аппаратно-программным модулем RCM

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Нагрузка на центральный процессор, %									
0,03	0,05	0,07	0,02	0,04	0,06	0,07	0,09	0,04	0,02
Загруженность SSD, %									
0,02	0,03	0,04	0,01	0,03	0,03	0,05	0,05	0,06	0,05
Потребление ОЗУ, %									
0,11	0,10	0,07	0,07	0,07	0,06	0,13	0,08	0,05	0,03

Таким образом, среднемесячная нагрузка на физические ресурсы, следующая:

- центральный процессор: 0,05 %;
- оперативная память: 0,07 %;
- твердотельный накопитель (SSD): 0,04 %.

4.3.2. Общая нагрузка ИмНС, РНС и MySQL при обработке больших разнородных и темпоральных данных на физические ресурсы вычислительного кластера

Нагрузка на физические ресурсы физического сервера, при отсутствии каких-либо внешних факторов (обычный режим), представлена в таблицах 4.17–4.19.

Таблица 4.17 – Загруженность CPU в течение одного месяца

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Количество обрабатываемой информации, Мбит/с									
40,00	100,00	160,00	220,00	280,00	340,00	400,00	460,00	520,00	580,00
Количество потоков, шт./с									
25,00	40,00	55,00	73,33	90,00	105,00	120,00	135,00	150,00	165,00
Средняя нагрузка на CPU при старте, %									
0,50	0,80	1,10	1,40	1,70	2,00	2,30	2,60	2,90	3,50
Средняя нагрузка на CPU при рестарте, %									
0,51	0,81	1,11	1,41	1,71	2,01	2,31	2,61	2,91	3,51
Средняя нагрузка на CPU при обработке информации, %									

0,75	1,20	1,65	2,10	2,55	3,00	3,45	3,90	4,35	4,60
Входящие сетевые пакеты, млн. шт./с									
0,40	1,00	1,60	2,20	2,80	3,40	4,00	4,60	5,20	5,80

Таблица 4.18 – Нагрузка на твердотельный накопитель (SSD) в течение одного месяца

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Количество обрабатываемой информации, Мбит/с									
40,00	100,00	160,00	220,00	280,00	340,00	400,00	460,00	520,00	580,00
Количество потоков, шт./с									
25,00	40,00	55,00	73,33	90,00	105,00	120,00	135,00	150,00	165,00
Средняя нагрузка на SSD при старте, %									
0,10	0,25	0,40	0,55	0,70	0,85	1,00	1,15	1,30	1,45
Средняя нагрузка на SSD при рестарте, %									
0,11	0,26	0,41	0,56	0,71	0,86	1,01	1,16	1,31	1,46
Средняя нагрузка на SSD при обработке информации, %									
0,20	0,50	0,80	1,10	1,40	1,70	2,00	2,30	2,60	2,90
Входящие сетевые пакеты, млн. шт./с									
0,40	1,00	1,60	2,20	2,80	3,40	4,00	4,60	5,20	5,80

Таблица 4.19 – Потребление ресурсов ОЗУ в течение одного месяца

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Количество обрабатываемой информации, Мбит/с									
40,00	100,00	160,00	220,00	280,00	340,00	400,00	460,00	520,00	580,00
Количество потоков, шт./с									
25,00	40,00	55,00	73,33	90,00	105,00	120,00	135,00	150,00	165,00
Среднее потребление на ОЗУ при старте, %									
0,18	0,48	0,72	0,96	1,20	1,44	1,68	1,92	2,16	2,40
Среднее потребление ОЗУ при рестарте, %									
0,19	0,49	0,73	0,97	1,21	1,45	1,69	1,93	2,17	2,41
Среднее потребление ОЗУ при обработке информации, %									
0,96	1,30	1,74	2,16	2,77	3,43	3,97	4,63	5,86	0,96
Входящие сетевые пакеты, млн. шт./с									
0,40	1,00	1,60	2,20	2,80	3,40	4,00	4,60	5,20	5,80

Средняя нагрузка (из таблиц 4.17–4.19):

- центральный процессор: 1,01%;
- SSD-накопитель: 0,53%;
- оперативная память: 0,67%.

Таким образом, по результатам тестирования (таблицы 4.17–4.19) было выявлено, что общая нагрузка на вычислительные ресурсы физического сервера ИмНС, РНС и MySQL является достаточно низкой и предоставляет возможность широкомасштабной обработки данных с дальнейшей стабилизацией серверной рабочей среды.

4.4. Итоговое тестирование и оценка эффективности прототипа СППР

Тестирование проводилось в течение одного календарного месяца: в целях компактности и наглядности брались средние результаты за три дня и вносились в таблицы. Также для более эффективного отслеживания динамики загруженности ресурсов графики строились на основе ежедневных данных.

Получение и анализ результатов нагрузки прототипом СППР (обычный режим и активный режим) представлены в таблицах 4.20–4.21.

Таблица 4.20 – Тестирование СППР в обычном режиме

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Средняя нагрузка на CPU, %									
2,52	2,73	2,28	2,20	2,79	2,62	2,34	2,16	2,22	2,54
Среднее потребление ОЗУ, %									
1,26	1,55	1,06	1,47	1,50	1,71	1,94	1,16	1,56	1,40
Средняя нагрузка на SSD, %									
1,17	1,15	1,19	1,56	1,36	1,83	1,59	1,48	1,11	1,05

Среднемесячная нагрузка на физические ресурсы составляет:

- центральный процессор: 2,44 %;
- оперативная память: 1,46 %;
- SSD: 1,35 %.

Таблица 4.21 – Тестирование прототипа СППР в активном режиме

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Средняя нагрузка на CPU, %									
4,7	4,44	4,07	4,50	4,76	4,50	4,45	4,32	4,07	4,06
Среднее потребление ОЗУ, %									
5,09	5,13	5,52	5,70	5,07	5,06	5,67	5,52	5,14	5,43
Средняя нагрузка на SSD, %									
6,38	6,18	6,43	6,03	6,37	6,37	6,45	6,53	6,85	6,16

Средний объем сетевого потока данных в одних календарных сутках приравнивается к 19 Гбит/с. Среднесуточное количество сетевых пакетов приравнивается к 10000000 в секунду. Отклонений в работоспособности рабочей среды физического сервера и кластера нет. Среднемесячные результаты нагрузки:

- центральный процессор: 4,38 %;
- оперативная память: 5,33 %;
- SSD: 6,15 %.

Сравнение нагрузки представлено в таблицах 4.22–4.24.

Таблица 4.22 – Сравнение нагрузки прототипа СППР в обычном и активном режимах

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Средняя нагрузка на CPU (обычный режим), %									
2,52	2,73	2,28	2,20	2,79	2,62	2,34	2,16	2,22	2,54
Средняя нагрузка на CPU (активный режим), %									
4,70	4,44	4,07	4,50	4,76	4,50	4,45	4,32	4,07	4,06
Удельный вес снижения нагрузки, %									
46,38	38,51	43,98	51,11	41,38	41,77	47,41	50,00	45,45	37,43

Таблица 4.23 – Сравнение потребления ОЗУ прототипа СППР в обычном и активном режимах

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Среднее потребление ОЗУ (обычный режим), %									

1,26	1,55	1,06	1,47	1,50	1,71	1,94	1,16	1,56	1,40
Среднее потребление ОЗУ (активный режим), %									
5,09	5,13	5,52	5,70	5,07	5,06	5,67	5,52	5,14	5,43
Удельный вес снижения нагрузки, %									
75,24	69,78	80,79	74,24	70,41	66,20	65,78	78,98	69,65	74,21

Таблица 4.24 – Сравнение нагрузки на SSD прототипа СППР в обычном и активном режимах

День									
1–3	4–6	7–9	10–12	13–15	16–18	19–21	22–24	25–27	28–30
Средняя нагрузка на SSD (обычный режим), %									
1,17	1,15	1,19	1,56	1,36	1,83	1,59	1,48	1,11	1,05
Средняя нагрузка на SSD (активный режим), %									
6,38	6,18	6,43	6,03	6,37	6,37	6,45	6,53	6,85	6,16
Удельный вес снижения нагрузки, %									
81,66	81,39	81,49	74,12	78,65	71,27	75,34	77,33	83,79	82,95

Исходя из данных таблиц 4.22–4.24 необходимо отметить, что замечено незначительное повышение нагрузки в активном режиме, связанное с повышенной обработкой сетевых пакетов, временным сохранением данных и их отправкой на веб-сервер.

Сравнение среднемесячной нагрузки представлено в таблице 4.25.

Таблица 4.25 – Среднемесячная нагрузка на физические ресурсы сервера

Название	Средняя нагрузка на CPU, %	Среднее потребление ОЗУ, %	Средняя нагрузка на SSD, %
Режим простоя	2,44	1,46	1,35
Активный режим	4,38	5,33	6,16

Таким образом, согласно таблицам 4.22–4.24 удельный вес снижения нагрузки является следующим:

- центральный процессор: 44,34%;
- оперативная память: 72,52%;

– твердотельный накопитель (SSD): 78,80%.

Фиксирование большого количества входящего сетевого потока данных (95 Гбит/с) разработанного прототипа СППР (за счет ИмНС) представлено на рисунке 4.4.

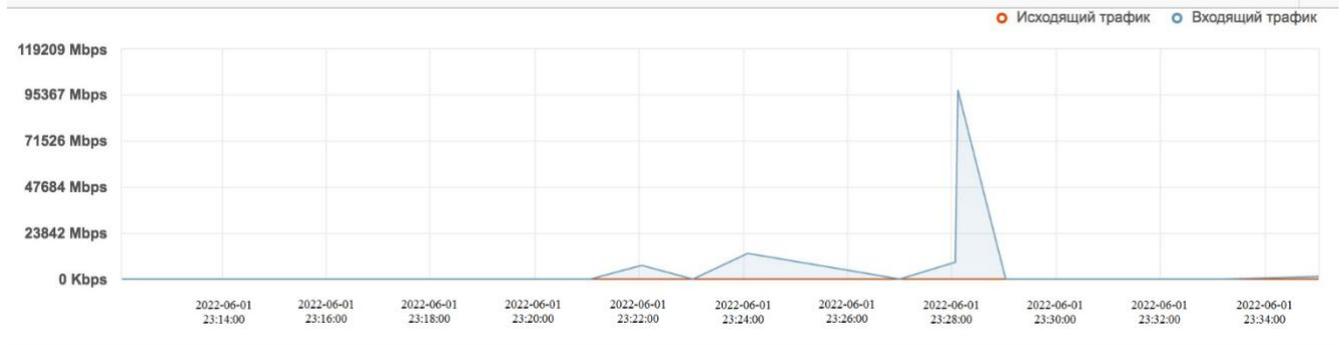


Рисунок 4.4 – Фиксирование аномалий в сетевом потоке данных

4.4.1. Показатель относительной эффективности

При оценке эффективности прототипа СППР важен и показатель относительной эффективности предлагаемого метода прогнозирования. В связи с тем, что известна средняя ошибка (в кВт*ч) по всем методам прогнозирования (таблицы 4.17–4.19), энергетическая организация в целях обеспечения устойчивой работы энергосистемы должна зарезервировать дополнительную мощность энергосистем, размер которой можно уточнить повышением точности прогнозирования. Следовательно, необходимо вычислить относительную эффективность Epo_{ji}^k (таблица 4.26) каждого рассматриваемого метода в диссертации по следующему соотношению:

$$Epo_{ji}^k = \left(1 - \frac{Ep_{cp}^k}{Em_{cp}^k} \right) \cdot 100\% , \quad (4.10)$$

где Ep_{cp}^k – средняя ошибка (кВт*ч) предлагаемого метода прогнозирования, Em_{cp}^k – средняя ошибка (кВт*ч) метода, с которым происходит сравнение (кВт*ч).

Таблица 4.26 – Расчет относительной эффективности модели рекуррентной ИНС

Прогнозная модель	Общее среднее значение ошибок по всем дням, кВт*ч	Относительная эффективность, %
Модель рекуррентной ИНС (раздел 2.3)	457,20	X
ARIMA	2187,50	-79%
SVR	2113,60	-78,4%
ARIMAX-GARCH	1784,50	-74,4%
MLR	2341,90	-80,5%
Однослойный персептрон	2983,50	-84,7%
Многослойный персептрон	1571,00	-70,9%
Сверточная	742,70	-38,4%
Классическая рекуррентная	1043,20	-56,2%

В таблице 4.26 представлено сравнение относительной эффективности согласно формуле (4.10) разработанной модели рекуррентной ИНС с другими моделями прогнозирования потребления электроэнергии. Таблица может быть дополнена строками с реальными данными, используемыми в организации для расчета прогнозных значений потребления электроэнергии и данными об ошибке такого прогноза.

Относительная эффективность (таблица 4.26), представленная в виде прогнозной точности, предлагаемой модели рекуррентной ИНС (раздел 2.3) по сравнению с применением других методов прогнозирования варьируется в пределах от 38,4 до 84,7%, и может быть использована для расчета (коррекции) количества резервируемых мощностей, например, снижением (или увеличением) части резервируемой мощности, рассчитываемой на основе учета ошибки прогноза. Например, при сравнении со сверточной ИНС разработанная модель ИНС точнее на 38,4%. Следовательно, чем меньше необоснованный процент резервирования энергетических мощностей, тем меньше финансовых потерь, что является особенно полезным в текущих экономических реалиях.

Таким образом, реализованная рекуррентная искусственная нейронная сеть показала более точный и стабильный результат, что доказывает целесообразность применения предлагаемого решения в рамках прогнозирования потребления

электроэнергии. За счет данных прогнозов у предприятий появится возможность проводить энергосберегающие мероприятия, начиная от контроля рабочих режимов и сроков проверки приборов учета электрической энергии и заканчивая реконструкцией электрических сетей. Более того, за счет прогноза потребления электроэнергии также появляется возможность предсказывать и возможные аварии на электросетях. Это позволит энергетическим компаниям более эффективно принимать различные управленческие решения и использовать финансовые ресурсы, что положительно скажется не только на тарифных планах электроэнергии для физических и юридических лиц, но и предоставит возможность оптимизации бюджета энергетической компании.

Выводы по главе 4

1. Проанализирована эффективность принципиально нового решения для обработки больших разнородных темпоральных данных на основе модели импульсной (спайковой) нейронной сети: модифицируется модель спайкового нейрона, позволяющая реагировать на внешние воздействия, в нашем случае на изменение видов и типов внешнего сетевого потока данных, что положительно сказывается на качестве обучения ИмНС и фильтрации сетевого потока данных в рамках обработки больших объемов данных. Но при распределенной обработке больших объемов данных оптимизация, в том числе и снижение нагрузки также является важным аспектом при работе любой системы, в связи с чем было выявлено, что средняя нагрузка на центральный процессор варьируется от 0,3% до 6,9%. Достаточно низкая загруженность CPU позволяет запускать многочисленные ресурсоемкие процессы, как во время поступления большого количества сетевого потока данных, так и в режиме простоя без каких-либо потерь производительности. Средняя загруженность SSD-накопителей, объединенных в массив RAID-10 (1+0), колеблется от 0,4% до 4,0%. Столь небольшая нагрузка на твердотельный накопитель позволяет использовать SSD другими ресурсоемкими процессами.

Загрузка ОЗУ варьируется от 0,5 до 3,1%. Небольшое потребление ресурсов ОЗУ также говорит о целесообразности использования разработанной импульсной нейронной сети для обработки больших объемов сетевого потока данных. Таким образом, предложенный метод позволяет реализовать упреждающее реагирование на аномалии в сетевом потоке данных, а также отфильтровывать сетевой поток данных для формирования исходного датасета, что открывает возможность применения в различных системах управления для информационной поддержки принятия решений.

2. Проанализирована эффективность метода обработки больших разнородных данных РНС на примере прогнозирования временных рядов, основанный на модели ИНС второго поколения. Анализ эффективности показал, что ИНС является лучшим подходом при раннем (заблаговременном) прогнозировании. Эксперименты показали очевидное преимущество предлагаемого решения в рамках прогнозирования потребления электроэнергии: до 17,7% точнее относительно других ИНС и до 34% точнее относительно статистических/гибридных методов прогнозирования, а также до 95% быстрее относительно других методов обучения ИНС. Это достигается модификацией нейрона и метода обучения рекуррентной нейронной сети, а также качественным исходным датасетом, получаемым импульсной нейронной сетью (задача №1). Данные результаты доказывают эффективность ее применения в расчетах прогнозных значений количества потребляемой электроэнергии. При этом за счет данных прогнозов у предприятий появится возможность принимать управленческие решения: например, проводить энергосберегающие мероприятия, начиная от контроля рабочих режимов и сроков проверки приборов учета электрической энергии и заканчивая реконструкцией электрических сетей, а также резервированием генерирующих мощностей электроэнергетических систем. Это позволит более эффективно использовать финансовые ресурсы энергетическим компаниям, что положительно скажется не только на тарифных планах

электроэнергии для физических и юридических лиц, но и предоставит возможность оптимизации бюджета энергетической компании.

ЗАКЛЮЧЕНИЕ

В результате научного исследования были получены следующие результаты и выводы:

1. Разработаны метод и модель обработки сетевого потока данных с применением импульсной нейронной сети для информационной поддержки принятия управленческих решений. Разработанные метод и нейросетевая модель позволяют качественно и эффективно осуществлять информационную поддержку принятия управленческих решений при поступлении больших объемов сетевого потока данных через внешний сетевой интерфейс за счет формируемых на выходе правил фильтрации, низкой загруженности серверных ресурсов (CPU: 0,3-6,9%; RAM: 0,4-4,0%; SSD: 0,5-3,1%) и качественной обработки больших данных, что позволяет автоматически формировать исходный датасет для прогнозирования потребления электроэнергии рекуррентной нейронной сетью. Данный результат достигается в том числе и методом обучения импульсной нейронной сети, за счет которого ИмНС с «нуля» обучается за 20 минут.

2. Разработаны метод и модель прогнозирования многомерных временных рядов с применением модели рекуррентной нейронной сети для информационной поддержки принятия управленческих решений. Данный метод за счет модифицированного метода обучения и высокой точности прогнозных значений позволяет оперативно и эффективно осуществлять информационную поддержку принятия управленческих решений. При этом модифицированные нейрон и метод обучения позволили получить прогнозные значения высокой точности (до 17,7% точнее относительно других ИНС) и повысить скорость обучения на 83–95%, а эффективность метода и модели экспериментально подтверждена на реальном массиве ретроспективных данных, полученным в результате обработки сетевого потока данных, при этом средняя величина ошибки не превышала 7,8%.

3. Разработана четырехкаскадная структурная схема (контур) системы управления процессом информационной поддержки принятия управленческих

решений в электроэнергетике, включающая в себя два контура обучения, контур адаптации и контур регулирования на основе которых реализован комплекс задач информационной поддержки с целью повышения эффективности принятия управленческих решений на примере решения задач обработки больших объемов сетевого потока данных и расчета прогнозных значений потребления электроэнергии. Показаны архитектура прототипа СППР и алгоритмы работы ее модулей.

4. Разработан прототип СППР на основе нейросетевой обработки больших темпоральных данных и прогнозирования временных рядов. Разработанное программное обеспечение позволяет свести к минимуму риски финансовых убытков предприятия за счет принятия решений ЛППР и содержит в себе следующие функции: информационная поддержка принятия управленческих решений на основе автоматизированной обработки больших разнородных и темпоральных данных; на основе автоматизированной обработки больших данных и разработанных методов и моделей, представленных в главе 2, прогнозирование временных рядов и генерация правил фильтрации сетевого потока данных.

5. Выполнена оценка эффективности предложенных решений, реализованных в виде прототипа СППР на примере реального массива ретроспективных данных и сетевого потока данных организации электроэнергетики, показывающая эффективность и целесообразность применения предлагаемой системы поддержки принятия решений. Средняя точность прогноза прототипа СППР (модель рекуррентной ИНС) составляет 92,2%, а относительная эффективность предлагаемого прототипа СППР прогнозирования точнее остальных методов на 38,4–84,7%. Следовательно, чем меньше необоснованный процент резервирования энергетических мощностей, тем меньше финансовых потерь, что является особенно полезным в текущих экономических реалиях.

Созданный в ходе работы прототип СППР на основе обработки больших данных и прогнозирования временных рядов имеет перспективы дальнейшего развития в рамках доработки математических моделей, увеличения

функциональности и полной реализации собственной библиотеки Deep Learning на языке программирования Java, а разработанные методы позволяют подстроить систему поддержки принятия решений под широкий класс задач, в которых необходимы автоматическое формирование датасетов и ранний расчет прогнозных значений.

СПИСОК УСЛОВНЫХ ОБОЗНАЧЕНИЙ И СОКРАЩЕНИЙ

ИмНС – импульсная нейронная сеть

ИНС – искусственная нейронная сеть

ЛПР – лицо, принимающее решение

ПУЭ – правила устройства электроустановок

РНС – рекуррентная нейронная сеть

СППР – система поддержки принятия решений

ANN – Artificial Neural Network (Искусственная нейронная сеть)

ARIMA – Autoregressive Integrated Moving Average (Интегрированная модель авторегрессии – скользящего среднего)

ARIMAX-GARCH – AutoRegressive Integrated Moving Average eXtended – Generalized AutoRegressive Conditional Heteroscedasticity (Расширенная авторегрессионная интегрированная скользящая средняя – обобщенная авторегрессионная условная гетероскедастичность)

CFNDGM – Conformable Fractional Non-Homogeneous Discrete Grey Model (Согласованная дробно-неоднородная дискретная модель серого цвета)

DLN – Deep Learning Networks (Сети глубокого обучения)

EMD-SVR-PSO-AR-GARCH – Empirical Mode Decomposition – Support Vector Regression – Particle Swarm Optimization – ARMA-Generalized AutoRegressive Conditional Heteroscedasticity (Декомпозиция эмпирического режима – Регрессия опорного вектора – Оптимизация роя частиц – ARMA-Обобщенная авторегрессионная условная гетероскедастичность)

GARCH – Generalized Autoregressive Conditional Heteroscedastic model (Модель авторегрессионной условной гетероскедастичности)

GRU – Gated Recurrent Units (Управляемый рекуррентный блок)

HES – Holt-winters Exponential Smoothing (Экспоненциальное сглаживание Холта-Уинтерса)

LSTM – Long Short-Term Memory (Длинная цепь элементов краткосрочной памяти)

- MLN – Multi-Layer Neural Network (Многослойная нейронная сеть)
- MLR – Multiple Linear Regression (Множественная линейная регрессия)
- OSI – Open Systems Interconnection (The Open Systems Interconnection model – Модель взаимосвязи открытых систем)
- PNN – Pulsed Neural Network (Импульсная нейронная сеть)
- RAID – redundant array of independent disks (Резервированный массив независимых дисков)
- RMSE – Root-Mean-Square Error (Среднеквадратичная ошибка)
- STDP – Spike-timing-dependent plasticity (Пластичность, зависящая от времени всплеска)
- SVM – Support Vector Machine (Метод опорных векторов)
- SVR – Support Vector Regression (Регрессия опорного вектора)
- TCN – Temporal Convolutional Network (Временная сверточная сеть)
- TRF – Time Response Function (Функция времени отклика)
- VECM – Vector Error Correction Models (Векторные модели исправления ошибок)

СПИСОК ЛИТЕРАТУРЫ

1. Moreira A.M. A two-level formal model for Big Data processing programs / J.B. Neto, M.M. Anamaria, G. Vargas-Solar, M.A. Musicante // Science of Computer Programming. –2021. – Vol. 215. – pp. 102764.
2. Васильев В.И. Опыт разработки многофакторной нейросетевой прогнозной модели потребления электроэнергии в энергосистеме ОАО «Башкирэнерго» / В.И. Васильев, С.С. Валеев, К.Р. Виноградова, С.А. Горбатков, И.З. Шахмаев // Нейрокомпьютеры: разработка, применение. – 2007. – №10. – С. 36–41.
3. Ерошенко С.А. Опыт внедрения моделей краткосрочного прогнозирования выработки солнечных электростанций / С.А. Ерошенко, Е.С. Кочнева, П.А. Крючков, А.И. Хпльясмаа // Энергоэксперт. – 2018. – №2(66). – С. 64–68.
4. Нургаянова О.С. Обучение нейронной сети для прогнозирования свойств никелевых сплавов на основе генетического алгоритма / О.С. Нургаянова, Н.И. Юсупова // Труды института системного анализа РАН. – 2019. – Т. 69. № 4. – С. 22–28.
5. Черняховская Л.Р. Эвергетика как методологическая основа управления выявлением дефектов на предпроектной стадии жизненного цикла систем обработки данных / В. Е. Гвоздев, Л. Р. Черняховская, Д. В. Блинова // Онтология проектирования. – 2018. – Т. 8. №1(27). – С. 22–29.
6. Брекоткина Е.С. Информационная поддержка управления уязвимостью сложных распределенных систем на основе обработки пространственной информации / Е. С. Брекоткина, М. Б. Гузаиров, С. В. Павлов, А. С. Павлов, О. И. Христодуло // Моделирование, оптимизация и информационные технологии. – 2020. – Т. 8. № 2(29). – С. 1–12.
7. Xu. Y. Feature data processing: Making medical data fit deep neural networks / Y. Xu, Z. Liu, Y. Li., H. Hou., Y. Zhao, W. Guo, L. Cui // Future Generation Computer Systems. – 2020. – Vol. 109. – pp. 149–157.

8. Сотников А. И. Анализ временных рядов с использованием алгоритмов BIG DATA / А. И. Сотников // Информационно-технологический вестник. – 2015. – Т. 18. № 2. – С. 156–177.

9. J. Gokulraj Data consistency matrix-based data processing model for efficient data storage in wireless sensor networks / J. Gokulraj, J. Senthilkumar, Y. Suresh, V. Mohanraj // Computer Communications. – 2020. – Vol. 151. № 1. – pp. 172–182.

10. Abrar S. Perturbation of deep autoencoder weights for model compression and classification of tabular data / S. Abrar, M. D. Samad // Neural Networks. – 2022. – Vol. 156. – pp. 160–169.

11. Серебряная Л. В. Методы построения искусственных нейронных сетей для классификации данных / Л. В. Серебряная // Цифровая трансформация. – 2022. – Т. 28. № 1. – С. 20–26.

12. Alexeev A. Using a Fully Connected Convolutional Network to Detect Objects in Images / A. Alexeev, Y. Matveev, G. Kukharev // IEEE. – 2018. – С. 141–146.

13. Pascanu R. On the difficulty of training recurrent neural networks / R. Pascanu, T. Mikolov, B. Youshua // Proceedings of the 30th International Conference on Machine Learning, ICML 2013. – 2013. – pp. 1310–1318.

14. Lwin T. Parameter estimation in first-order autoregressive model for statistical process monitoring in the presence of data autocorrelation / T. Lwin // Journal of Statistical Planning and Inference. – 2011. – Vol. 141. № 8. – pp. 2556–2575.

15. Lundahl B. A statistical model of routine process monitoring variates and conformance assessment for radiation processing / B. Lundahl // Radiation Physics and Chemistry. – 2022. – pp. 110640.

16. Offermans T. Automatically optimizing dynamic synchronization of individual industrial process variables for statistical modelling / T. Offermans, E. Szymanska, G. H. Kollenburg, L.M. Buydens, J. J. Jansen // Computers & Chemical Engineering. – 2022. – Vol. 152. – pp. 117–147.

17. Zhichao L. An ensemble framework based on multivariate statistical analysis for process monitoring / L. Zhichao, L. Titan, Y. Xuefeng // *Expert Systems with Applications*. – 2022. – pp. 117732.

18. Rasay H. An integrated Maintenance and Statistical Process Control Model for a Deteriorating Production Process / H. Rasay, S. Taghipour, M. Sharifi // *Reliability Engineering & System Safety*. – 2022. – Vol. 228. – pp. 108774.

19. Бартоломей П.И. Разработка методов и алгоритмов обнаружения плохих данных и повышения качества телеметрии в задачах управления энергосистемами / П.И. Бартоломей, Е.Н. Бегалова // *Известия высших учебных заведений. Проблемы энергетики*. – 2008. – №11-12. – С. 40–49.

20. Vasilyev V.I. Bank transaction text label mining algorithms / A.S. Startseva, A.M. Vulfin, V.I. Vasilyev, A.V. Nikonov, A.D. Kirillova // *Информационные технологии и нанотехнологии (ИТНТ-2020): Сборник трудов по материалам VI Международной конференции и молодежной школы – 2020*. – Т. 4. – С. 445–454.

21. Тагирова К.Ф. Разработка обобщенного алгоритма обработки данных различной физической природы на основе системного анализа / К.Ф. Тагирова, В.А. Кощев, А.М. Вульфин, К.В. Миронов // *Информационные технологии и системы: Труды Седьмой Всероссийской научной конференции с международным участием*. – 2019. – С. 26–31.

22. Шмелев О.С. Методы обработки больших массивов данных / О.С. Шмелев, Е.М. Гриценко // *Лесной и химический комплексы – проблемы и решения: Сборник материалов по итогам Всероссийской научно-практической конференции*. – 2022. – С. 515–519.

23. Полтавцева М.А. Построение адаптивной системы сбора и подготовки данных при мониторинге безопасности / М.А. Полтавцева, Д.П. Зегжда // *Проблемы информационной безопасности. Компьютерные системы*. – 2020. – №2. – С. 92–101.

24. Тагирова К.Ф. Разработка обобщенного алгоритма обработки данных различной физической природы на основе системного анализа / К.Ф. Тагирова,

В.А. Кощев, А.М. Вульфин, К.В. Миронов // Информационные технологии и системы: Труды Седьмой Всероссийской научной конференции с международным участием. – 2019. – С. 26–31.

25. Четырбок П.В. Большие данные и глубокое машинное обучение в искусственных нейронных сетях / П.В. Четырбок, Р.П. Ганусяк // Информационные системы и технологии в моделировании и управлении: сборник трудов V Международной научно-практической конференции. – 2020. – С. 100–103.

26. Devi B.S. Symmetric Measure of Network Traffic using Packet Ratio and Packet Symmetry / B.S. Devi, K. Dudeja, A.V. John, K. Marcin, T. Subbulakshmi // Procedia Computer Science. – 2019. – Vol. 165. – pp. 112–118.

27. Peng T. Urban fire emergency management based on big data intelligent processing system and Internet of Things / T. Peng, K. Wang // Optik. – 2022. – Vol. 273. – pp. 170433.

28. Грицай А.С. Исследование факторов, влияющих на потребление электроэнергии коммерческим предприятием / А.В. Блохин, А.С. Грицай, А.Ю. Горшенин // Математические структуры и моделирование. – 2022. – № 3(63). – С. 39-47.

29. Затонский А.В. Выбор модели прогнозирования отпуска тепловой энергии / А.В. Затонский, Л.Г. Тушакова // Теплоэнергетика. – 2021. – №3. – С. 89–98.

30. Вялкова С.А. Прогнозирование суточных графиков активной мощности мегаполиса с учетом прогнозных данных естественной освещенности / С.А. Вялкова, И.И. Надтока // Известия высших учебных заведений. Электромеханика. – 2020. – Т. 63. №5. – С. 67–71.

31. Корнейчук, О.С. Прогнозирование объемов потребления электроэнергии на основе статистических данных / О.С. Корнейчук, И.В. Фёдоров // Прикладная математика и фундаментальная информатика. – 2018. – Т. 5., № 1. – С. 13–20.

32. Филимонов, Н.Б. Идентификация динамических систем на основе обработки экспериментальных данных методом генетического программирования

/ Л. Чжан, Н.Б. Филимонов // Журнал передовых исследований в области естествознания. – 2023. – №18. – С. 4–12.

33. Juyong L., Youngsang C. National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model? / L. Juyong, C. Youngsang // Energy. – 2022. – No 239-D. – pp. 122366.

34. Hadjout D., et.al. Electricity consumption forecasting based on ensemble deep learning with application to the Algerian market / D. Hadjout // Energy. – 2022. – No 243. – pp. 123060.

35. Hai-Bao C., Ling-Ling P., Yu-Feng Z. Forecasting seasonal variations in electricity consumption and electricity usage efficiency of industrial sectors using a grey modeling approach / C. Hai-Bao, P. Ling-Ling, Z. Yu-Feng // Energy. – 2021. – No 222. – pp. 119952.

36. Nishant J., et.al. Electricity load forecasting and feature extraction in smart grid using neural networks / J. Nishant // Computers & Electrical Engineering. – 2021. – No 96-A. – pp. 107479.

37. Xu R. Z. An integrated approach based on virtual data augmentation and deep neural networks modeling for VFA production prediction in anaerobic fermentation process / R. Z. Xu, J. S. Cao, Y. Wu, S. N. Wang, J. Y. Luo, X. Chen, F. Fang // Water Research. – 2020. – Vol 184. – pp. 116103.

38. Wu Q. Data processing platform design and algorithm research of wearable sports physiological parameters detection based on medical internet of things / Q. Wu, P. Tang, M. Yang // Measurement. – Vol. 165. № 1. – Elsevier, 2020. – pp. 108172.

39. Jinghua L., Yichen L., Shanyang W. Long-term electricity consumption forecasting method based on system dynamics under the carbon-neutral target/ L. Jinghua, L. Yichen, W. Shanyang // Energy. – 2022. – No 244–A. – pp. 122572.

40. Ying S., et.al. Prediction method of electricity stealing behavior based on multi-dimensional features and BP neural network / S. Yimg // Energy Reports. – 2022. – No. 8–4. – pp. 523–531.

41. Аверченкова Е.Э. Принятие решений в системе управления реализацией национальных проектов в регионах РФ на основе математических и алгоритмических методов системного анализа / Е.Э. Аверченкова // Научно-технический вестник Поволжья. – 2022. – №12. – С. 86–89.

42. Антонов В. В. Особенности извлечения данных из социальных сетей / В.В. Антонов, Л.Е. Родионова, Н.Р. Калимуллин, Я.С. Вояковская // Охрана, безопасность, связь. – 2021. – № 6–2. – С. 168–175.

43. Вересников Г.С. Оптимизационные модели параметрического синтеза проектных решений в условиях неопределенности параметров / Г.С. Вересников // Управление большими системами: сборник трудов. – 2020. – № 85. – С. 238–257.

44. Гвоздев В. Е. Информационная поддержка принятия решений по устранению дефектов, выявленных при эксплуатации систем обработки данных / В. Е. Гвоздев, Д. В. Блинова, К. В. Кириллов // Информационные технологии. – 2019. – Т. 25. № 3. – С. 165–173.

45. Гончаренко В.И. Алгоритм нечеткой логики принятия решения при определении допустимой длительности обслуживания заявки в динамической обстановке / В.И. Гончаренко, Г.Н. Лебедев, Д.А. Михайлин, Н.Г. Журавлева // Цифровая обработка сигналов и её применение (ЦОС-2021). – 2021. – С. 179–183.

46. Васильев, В.И. Система поддержки принятия решений по обеспечению информационной безопасности автоматизированной системы управления технологическими процессами / В. И. Васильев, В. Е. Гвоздев, М. Б. Гузаиров, А. Д. Кириллова // Информация и безопасность. – 2017. – Т. 20, № 4. – С. 618–623.

47. Ильясов Б. Г. Совершенствование потребительских свойств систем обработки данных по результатам их эксплуатации / Б.Г. Ильясов, В.Е. Гвоздев, Д.В. Блинова, Р.А. Мунасыпов // Информационные технологии интеллектуальной поддержки принятия решений. Труды V Всероссийской конференции (с приглашением зарубежных ученых). – 2017. – С. 50–53.

48. Кулида Е.Л. Поддержка принятия управленческих решений в сфере жилищно-коммунального хозяйства // Е.Л. Кулида, Л.А. Панкова //

Информационные технологии и вычислительные системы. – 2009. – № 4. – С. 93–103.

49. Маторин С.И. Системно-объектный детерминантный анализ. Построение генетической и партитивной классификаций предметной области / С.И. Маторин, В.В. Михелев // Искусственный интеллект и принятие решений. – 2022. – №1. – С. 26–34.

50. Ефремова, О.А. Системный анализ проблемы организации поддержки принятия решений органов исполнительной власти с использованием пространственной информации / О. А. Ефремова, С. В. Павлов // Естественные и технические науки. – 2016. – № 12(102). – С. 255–259.

51. Сагитова В.В. Применение метода экспертных оценок для автоматизации аудита информационных систем персональных данных / В. В. Сагитова, В. И. Васильев // Вестник УГАТУ. – 2017. – Т. 21. № 3 (77). – С. 105–112.

52. Тиханычев О.В. Постановка задачи обеспечения автоматизации полного цикла поддержки принятия решения / О.В. Тиханычев // Прикладная информатика. – 2021. – Т. 16. №6(96). – С. 103–112.

53. Yusupova, N. I. Analysis of Object State Data Based on Integration of Artificial Intelligence Technologies / N. I. Yusupova, G. R. Shakhmametova // Proceedings – ICOECS 2020: 2020 International Conference on Electrotechnical Complexes and Systems, Ufa, 27–30 октября 2020 года. – 2020. – pp. 9278414.

54. Chinyere N.O. Adoption and utilization of medical decision support systems in the diagnosis of febrile Diseases: A systematic literature review / N.O. Chinyere, K. Nayab, E. Victory, A. Kingsley, N. Chukwudi, S. Humphrey, A. Christie, O. Joseph, B. Luis, F.M. Yzoka // Expert Systems with Applications. – 2023. – Vol. 220. – pp. 119638.

55. Sutton R.T. An overview of clinical decision support systems: benefits, risks, and strategies for success / R.T. Sutton, D. Pincock, D.C. Baumgart, D.C. Sadowski, R.N. Fedorak, K.I. Kroeker // npj Digital Medicine. – 2020. – Vol. 3. No 17. – pp. 1–11.

56. Концепция «Цифровая трансформация 2030» – URL: https://www.rossetivolga.ru/i/files/2019/2/7/kontseptsiya_tsifrovaya_transformatsiya_2030.pdf (Дата обращения 20.11.2022).

57. Mofrad F. B. DenseNet-based Transfer Learning for LV Shape Classification: Introducing a Novel Information Fusion and Data Augmentation using Statistical Shape/Color Modeling / F. B. Mofrad, G. Valizadeh // Expert Systems with Applications. – 2022. – Vol. 213. Part C. – pp. 119261.

58. НЭБ «eLIBRARY» – <https://elibrary.ru/> (Дата обращения 19.12.2023)

59. Мифы и реальность "устойчивого развития" / Г. С. Розенберг, С. А. Черникова, Г. П. Краснощеков [и др.] // Проблемы прогнозирования. – 2000. – № 2. – С. 131-154.

60. Троицкий, А. А. Энергетическая стратегия - важнейший фактор социально-экономического развития России / А. А. Троицкий // Теплоэнергетика. – 2001. – № 7. – С. 2-9.

61. Васильев, И. Е. Математическая модель расчёта и прогнозирования удельного расхода электроэнергии при производстве водорода / И. Е. Васильев, Р. В. Ключев // Известия высших учебных заведений. Электромеханика. – 2002. – № 3. – С. 59-62.

62. Шакаев, С. М. Потери электроэнергии в электрических сетях ОАО "Татэнерго" / С. М. Шакаев, А. И. Федотов, Р. Н. Камалиев // Известия высших учебных заведений. Проблемы энергетики. – 2003. – № 3-4. – С. 102-107.

63. Попова, Е. Б. Прогнозирование рационального электропотребления молочно-товарных ферм / Е. Б. Попова // Инженерно-техническое обеспечение АПК. Реферативный журнал. – 2004. – № 3. – С. 872.

64. Алескеров, Ф. Т. Производство, легальное и нелегальное потребление электроэнергии (динамический анализ) / Ф. Т. Алескеров, Б. М. Шит // Проблемы управления. – 2005. – № 2. – С. 63-69.

65. Карякин, А. Л. Прогнозирование потребления электроэнергии цехом агломерации на основе метода искусственных нейронных сетей / А. Л. Карякин, К. Д. Белов, Е. А. Дегтярев // Известия высших учебных заведений. Горный журнал. – 2008. – № 8. – С. 126-127.

66. Кумаритов, А. М. Прогнозирование потребления энергоресурсов в учреждениях бюджетной сферы / А. М. Кумаритов, И. С. Портнов // Вестник Воронежского государственного технического университета. – 2009. – Т. 5, № 10. – С. 201-204.

67. Коробкова, Е. А. Оценка перспектив развития угледобычи на основе прогнозных экономических показателей / Е. А. Коробкова, Б. Л. Тальгамер // Вестник Иркутского государственного технического университета. – 2010. – № 5(45). – С. 70-76.

68. Иващенко, В. А. Прогнозирование электропотребления промышленных предприятий на основе статистических методов и искусственных нейронных сетей / В. А. Иващенко, М. В. Колоколов, Д. А. Васильев // Вестник Саратовского государственного технического университета. – 2010. – Т. 2, № 1(45). – С. 110-115.

69. Копцев, Л. А. Нормирование и прогнозирование потребления электроэнергии на промышленном предприятии / Л. А. Копцев, А. Л. Копцев // Промышленная энергетика. – 2011. – № 1. – С. 18-23.

70. Гофман, А. В. Повышение точности краткосрочного и оперативного прогнозирования электропотребления энергосистемы с применением искусственной нейронной сети / А. В. Гофман, А. С. Ведерников, Е. С. Ведерникова // Электрические станции. – 2012. – № 7(972). – С. 36-41.

71. Размер обучающей выборки и ее влияние на архитектуру искусственной нейронной сети в энергосистеме / В. З. Манусов, И. С. Макаров, С. А. Дмитриев [и др.] // Вестник Тамбовского университета. Серия: Естественные и технические науки. – 2013. – Т. 18, № 4-1. – С. 1417-1420.

72. Филатова, Е. С. Методика построения системы прогнозирования электропотребления на базе нейронечеткого подхода и вейвлет-теории / Е. С.

Филатова, Д. М. Филатов, А. Д. Стоцкая // Международный журнал прикладных и фундаментальных исследований. – 2014. – № 11-1. – С. 26-29.

73. Староверов, Б. А. Определение наиболее перспективных нейронных сетей и способов их обучения для прогнозирования электропотребления / Б. А. Староверов, Б. А. Гнатюк // Вестник Ивановского государственного энергетического университета. – 2015. – № 6. – С. 59-64.

74. Дякин, Н. В. Прогнозирование выработки и потребления электроэнергии в многоагентной системе / Н. В. Дякин // Практическая силовая электроника. – 2016. – № 2(62). – С. 53-56.

75. Иванов, В. В. Прогнозирование суточного потребления электроэнергии в Московском регионе с использованием искусственных нейронных сетей / В. В. Иванов, А. В. Крянев, Е. С. Осетров // Письма в журнал Физика элементарных частиц и атомного ядра. – 2017. – Т. 14, № 4(209). – С. 418-432.

76. Корнейчук, О. С. Прогнозирование объёмов потребления электроэнергии на основе статистических данных / О. С. Корнейчук, И. В. Федоров // Прикладная математика и фундаментальная информатика. – 2018. – Т. 5, № 1. – С. 13-20. – DOI 10.25206/2311-4908-2018-5-1-13-20.

77. Прогнозирование потребления электрической энергии дистанции электроснабжения с помощью искусственных нейронных сетей / В. Н. Галушко, В. А. Пацкевич, И. С. Евдасев, А. В. Дробов // Вестник Белорусского государственного университета транспорта: наука и транспорт. – 2019. – № 1(38). – С. 39-41.

78. Барбасова, Т. А. Метод многоуровневого прогнозирования и нормирования потребления электрической энергии предприятием / Т. А. Барбасова // Электротехнические системы и комплексы. – 2020. – № 4(49). – С. 49-54.

79. Рахмонов, И. У. Прогнозирование электропотребления промышленных предприятий с помощью модели авторегрессии проинтегрированного скользящего среднего / И. У. Рахмонов, Н. Н. у. Курбонов // Вестник Московского энергетического института. Вестник МЭИ. – 2021. – № 6. – С. 11-19.

80. Прогнозирование потребления электрической энергии промышленным предприятием с помощью методов машинного обучения / А. Д. Моргоева, И. Д. Моргоев, Р. В. Ключев, О. А. Гаврина // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2022. – Т. 333, № 7. – С. 115-125.

81. Полуянович, Н. К. Нейросетевое прогнозирование энергопотребления как основополагающий принцип устойчивой энергетической системы / Н. К. Полуянович, М. Н. Дубяго // Известия Ошского технологического университета. – 2023. – № 2–1. – С. 187–193.

82. Zongwei M. A review of statistical methods used for developing large-scale and long-term PM_{2.5} models from satellite data / M. Zongwei, D. Sagnik, S. Christopher, R. Liu, J. Bi, P. Balyan, Y. Liu // Remote Sensing of Environment. – 2022. – Vol. 269. – pp. 112827.

83. Yamaguchi D. A grey-based rough approximation model for interval data processing / D. Yamaguchi, G. D. Li, M. Nagai // Information Sciences. – 2007. – Vol. 177. № 21. – pp. 4727–4744.

84. Liu Y. A conformable fractional unbiased grey model with a flexible structure and it's application in hydroelectricity consumption prediction / Y. Liu, Y. Yang, F. Pan, D. Xue // Journal of Cleaner Production. – 2022. – Vol. 367. – pp. 133029.

85. Zeng B. Modeling, prediction and analysis of new energy vehicle sales in China using a variable-structure grey model / B. Zeng, L. Hui, C. Mao, Y. Wu // Expert Systems with Applications. – 2023. – Vol 213. Part A. № 1. – pp. 118879.

86. Zeng J. Forecasting the sustainable classified recycling of used lithium batteries by gray Graphical Evaluation and Review Technique / J. Zeng, L. Sifeng // Renewable Energy. – 2022. – Vol. 202. – pp. 602–612.

87. Xuemei L. A novel seasonal grey model for forecasting the quarterly natural gas production in China / L. Xuemei, G. Xinchang, L. Lina, C. Yun, B. Yang // Energy Reports. – 2022. – Vol. 8. – pp. 9142–9157.

88. Cai K. Using grey Gompertz model to explore the carbon emission and its peak in 16 provinces of China / K. Cai, W. Lifeng // *Energy and Buildings*. – 2022. – Vol. 277. – pp. 112545.

89. Yan S. Online public opinion prediction based on rolling fractional grey model with new information priority / S. Yan, Q. Su, Z. Ging, X. Zeng, E. Herrera-Viedma // *Information Fusion*. – 2023. – Vol. 91. – pp. 277–298.

90. He X. A novel structure adaptive new information priority discrete grey prediction model and its application in renewable energy generation forecasting / X. He, Y. Wang, Y. Zhang, X. MA, W. Wu, L. Zhang // *Applied Energy*. – 2022. – Vol. 325. – pp. 119854.

91. Zeng X. A novel matrixed nonlinear grey Bernoulli model for interval prediction of power generation / X. Zeng, C. Haoze, S. Yan, S. Lv // *Applied Mathematical Modelling*. – 2023. – Vol. 115. – pp. 221–236.

92. Dong L. Applicability of statistical and modeling methods in assessing the basin-scale grey water footprint / L. Dong, X. Zhuohang, C. Song, C. Zhang, X. Bai // *Ecological Indicators*. – 2022. – Vol. 143. – pp. 109345.

93. Parody L. Hybrid business process modeling for the optimization of outcome data / L. Parody, M. Gomez-Lopez, R. M. Gasca // *Information and Software Technology*. – 2016. – Vol. 70. – pp. 140–154.

94. Flaus J. M. Data flow modeling for batch and hybrid processes / J. M. Flaus, L. Thevenon // *ISA Transactions*. – 2003. – Vol. 42, № 3. – pp. 361–380.

95. Shah P. Deep neural network-based hybrid modeling and experimental validation for an industry-scale fermentation process: Identification of time-varying dependencies among parameters / P. Shah, M. Sheriff, M. Bangi, C. Kravaris, J. Sang-Kwon, C. Botre, J. Hirota // *Chemical Engineering Journal*. – 2022. – Vol. 441. № 1. – pp. 135643.

96. Wenli L. A Hybrid Data-driven Model for Geotechnical Reliability Analysis / L. Wenli, A. Li, F. Weili, P. Love, T. Hartmann, H. Luo // *Reliability Engineering & System Safety*. – 2023. – Vol. 231. – pp. 108985.

97. Huang W. Tool wear prediction in ultrasonic vibration-assisted drilling of CFRP: A hybrid data-driven physics model-based framework / W. Huang, X. Zhang, C. Wu, S. Cao, Q. Zhou // *Tribology International*. – 2022. – Vol. 174. – pp. 107755.

98. McKay A. A novel linear hybrid model predictive control design: application to a fed batch crystallization process / A. McKay, D. Ghosh, L. Zhu, L. Xi, P. Mhaskar // *Digital Chemical Engineering*. – 2022. – Vol. 3. – pp. 100033.

99. Shi T. The lending risk predicting of the folk informal financial organization from big data using the deep learning hybrid model / T. Shi, C. Li, H. Wanyan, Y. Xu, W. Zhang // *Finance Research Letters*. – 2022. – Vol. 50. – pp. 103212.

100. Mohan V.N. Hybrid machine learning-based model for solubilities prediction of various gases in deep eutectic solvent for rigorous process design of hydrogen purification / V. N. Mohan, H. M. Rehman, J. Haider, M. Abdul-Qyyum, G. Sang Choi, H. Lim // *Separation and Purification Technology*. – 2022. – Vol. 298. № 1. – pp. 121651.

101. Yuting L. Predicting ammonia nitrogen in surface water by a new attention-based deep learning hybrid model / L. Yuting, L. Ruying // *Environmental Research*. – 2023. – Vol. 216. Part 3. № 1. – pp. 114723.

102. De Coste M. The prediction of mid-winter and spring breakups of ice cover on Canadian rivers using a hybrid ontology-based and machine learning model / M. De Coste // *Environmental Modelling & Software*. – 2022. – Vol. 160. – pp. 105577.

103. Schmidhuber J. Deep Learning in Neural Networks: An Overview / J. Schmidhuber // *Neural Networks*. – 2015. – Vol. 61. – pp. 85–117.

104. Ptukhin A.A. Machine learning in the processing and analysis of texts / A.A. Ptukhin // *Languages in professional communication*. – 2019. – pp. 517–523.

105. Сукатова А. Ю. Применение искусственных нейронных сетей для прогнозирования шероховатости при чистовом и получистовом точении / Д. С. Баранов, Т. А. Дуюн // *Вестник белгородского государственного технологического университета им. В.Г. ШУХОВА*. – 2019. – № 7. – С. 128–134.

106. Дауб И. С. Обзор методов прогнозирования временных рядов с помощью искусственных нейронных сетей / И. С. Дауб // STUDNET. – 2020. – Т. 3. № 10. – С. 247.

107. Pimenov V. I. Interpretation of a trained neural network based on genetic algorithms / V. I. Pimenov, I. V. Pimenov // Information and control systems. – № 6(109). – 2020. – С. 12–20.

108. Сукатова А. Ю. Нейросетевые методы обработки информации в задачах прогноза климатических характеристик / А. Ю. Сукатова, // Наука и современность. – 2010. – С. 1–3.

109. Zhong X. C. A process-data-driven BP neural network model for predicting interval-valued fatigue life of metals / X. C. Zhong, R.K. Xie, S. H. Qin, K. S. Zhang // Engineering Fracture Mechanics. – 2022. – Vol. 256. Part A. – pp. 108918.

110. Zheng W. A data processing method for electronic tongue based on computational model of taste pathways and convolutional neural network / W. Zheng, Y. Shi, X. Xia, Y. Ying, H. Men // Measurement. – 2022. – Vol. 205. – pp. 112150.

111. Fugh J. N. Local approximate Gaussian process regression for data-driven constitutive models: development and comparison with neural networks / J. N. Fugh, M. Marino, N. Bouklas // Computer Methods in Applied Mechanics and Engineering. – 2021. – Vol. 388. – pp. 114217.

112. Koronaki E. D. A data-driven reduced-order model of nonlinear processes based on diffusion maps and artificial neural networks / E. D. Koronaki, A. M. Nikas, A. G. Boudouvis // Chemical Engineering Journal. – 2022. – Vol. 397. – pp. 125475.

113. Feng B. A pre-processing method for independent historical data in artificial neural network models: A case study in Laizhou Bay / B. Feng, R. Hu // Regional Studies in Marine Science. – 2020. – Vol. 35. – pp. 101238.

114. Menegozzo G. Automatic process modeling with time delay neural network based on low-level data / G. Menegozzo, D. Dall’Alba, A. Roberti, P. Fiorini // Procedia Manufacturing. – 2019. – Vol. 38. – pp. 125–132.

115. Xiao H. Fault diagnosis and prognosis of wastewater processes with incomplete data by the auto-associative neural networks and ARMA model / H. Xiao, D. Huang, Y. Pan, Y. Liu, K. Song // *Chemometrics and Intelligent Laboratory Systems*. – 2017. – Vol. 161. – pp. 96–107.

116. Ahmad Z. Combination of multiple neural networks using data fusion techniques for enhanced nonlinear process modelling / Z. Ahmad, J. Zhang // *Computers & Chemical Engineering*. – 2005. – Vol. 30. № 2. – pp. 295–308.

117. Sudha C. Accurate data aggregation created by neural network and data classification processed through machine learning in wireless sensor networks / C. Sudha, D. Suresh, A. Nagesh // *Theoretical Computer Science*. – 2022. – Vol. 925. pp. 25–36.

118. Yang L. Neural network stochastic differential equation models with applications to financial data forecasting / L. Yang, T. Gao, Y. Lu, J. Duan, T. Liu // *Applied Mathematical Modelling*. – 2023. – Vol. 115. – pp. 277–299.

119. Wang M. Data-driven strain–stress modelling of granular materials via temporal convolution neural network / M. Wang, T. Qu, S. Guan, T. Zhao, B. Liu, Y. T. Feng // *Computers and Geotechnics*. – 2022. – Vol. 152. – pp. 105049.

120. Araujo N. Multilayer perceptron neural network applied to TG dynamic data of biopolymer chitosan – A robust tool to study the kinetics of solid thermal decomposition / N. Araujo, R. Sebastiao, M. Freitas-Marques, W. Nova Mussel, M. Irene Yoshida, L. S. Virtuoso // *Thermal Science and Engineering Progress*. – 2022. – Vol. 36. – pp. 101490.

121. Napoli C. Exploiting Wavelet Recurrent Neural Networks for satellite telemetry data modeling, prediction and control / C. Napoli, G. De Magistris, C. Ciancarelli, F. Corallo, F. Russo, D. Nardi // *Expert Systems with Applications*. – 2022. – Vol. 206. – pp. 117831.

122. Arpad G., et.al. Forecasting Electricity Consumption and Production in Smart Homes through Statistical Methods / G. Arpad // *Sustainable Cities and Society*. – 2022. – No. 76. – pp. 103426.

123. Pınar C. Impact of the COVID-19 lockdowns on electricity and natural gas consumption in the different industrial zones and forecasting consumption amounts: Turkey case study / C. Pınar // International Journal of Electrical Power & Energy Systems. – 2022. – No. 134. – pp. 107369.

124. Rehan J. Hydroelectricity consumption forecast for Pakistan using ARIMA modeling and supply-demand analysis for the year 2030 / J. Rehan // Renewable Energy. – 2020. – No. 154. – pp. 1–10.

125. Pyramid Arima – <https://pypi.org/project/pyramid-arima/> (Дата обращения 26.06.2023).

126. Guo-Feng F., et.al. Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling / F. Guo-Feng // Utilities Policy. – 2021. – No. 73. – pp. 101294.

127. Ping J., et.al. A novel composite electricity demand forecasting framework by data processing and optimized support vector machine / J. Ping // Applied Energy. – 2020. – No. 260. – pp. 114243.

128. Guo-Feng F., et.al. Forecasting electricity consumption using a novel hybrid model / F. Guo-Feng // Sustainable Cities and Society. – 2020. – No 61. – pp. 102320.

129. Ashutosh K.D., et.al. Study and analysis of SARIMA and LSTM in forecasting time series data / K.D. Ashutosh // Sustainable Energy Technologies and Assessments. – 2021. – No 47. – pp. 101474.

130. Kianoosh G., et.al. A novel multi-time-scale modeling for electric power demand forecasting: From short-term to medium-term horizon / G. Kianoosh // Electric Power Systems Research. – 2017. – No 142. – pp. 58–73.

131. Alireza S., Mehdi Z. The electricity consumption forecast: Adopting a hybrid approach by deep learning and ARIMAX-GARCH models / S. Alireza, Z. Mehdi // Energy Reports. – 2022. – No 8. – pp. 7657–7679.

132. Weiheng J. Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption // J. Weiheng, X. Wu, Y. Gong, W. Yu, X. Zhong // Energy. – 2020. – Vol. 193. – pp. 116779.

133. Zhenyu Z., et.al. Load forecasting via Grey Model-Least Squares Support Vector Machine model and spatial-temporal distribution of electric consumption intensity / Z. Zhenyu // Energy. – 2022. – No 255. – pp. 124468.

134. Tao T. et.al. GM(1,1) based improved seasonal index model for monthly electricity consumption forecasting / T. Tao // Energy. – 2022. – No. 252. – pp. 124041.

135. Bin P. UFNGBM (1,1): A novel unbiased fractional grey Bernoulli model with Whale Optimization Algorithm and its application to electricity consumption forecasting in China / P. Bin // Energy Reports. – 2022. – No 7. – pp. 7405–7423.

136. Wenhao Z., Hailin L., Zhiwei Z. A novel seasonal fractional grey model for predicting electricity demand: A case study of Zhejiang in China / Z. Wenhao, L. Hailin, Z. Zhiwei // Mathematics and Computers in Simulation. – 2022. – No 200. – pp. 128–147.

137. Yitong L., Dingyu L., Yang Y. Two types of conformable fractional grey interval models and their applications in regional electricity consumption prediction / L. Yitong, L. Dingyu, Y. Yang // Chaos, Solitons & Fractals – 2021. – No 153-2. – pp. 111628.

138. Haoran Ma. Prediction of industrial power consumption in Jiangsu Province by regression model of time variable / Ma Haoran // Energy. – 2022. – No 239-B. – pp. 122093.

139. Jui-Sheng C., Duc-Son T. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders / C. Jui-Sheng, T. Duc-Son // Energy. – 2018. – No 165-B. – pp. 709–726.

140. Mustapha L.A., et.al. A Review on Deep Learning with Focus on Deep Recurrent Neural Network for Electricity Forecasting in Residential Building / L.A. Mustapha // Procedia Computer Science. – 2021. – No 193. – pp. 141–154.

141. Zhuang Z., Hainan C., Xiaowei L. Spatial granularity analysis on electricity consumption prediction using LSTM recurrent neural network / Z. Zhuang, C. Hainan, L. Xiaowei // Energy Procedia. – 2019. – No 158. – pp. 2713–2718.

142. Che L., et.al. A hybrid prediction model for residential electricity consumption using holt-winters and extreme learning machine / L. Che // *Applied Energy*. – 2020. – No 275. – pp. 115383.

143. Sulaiman, S.M. et.al. A novel hybrid short-term electricity forecasting technique for residential loads using Empirical Mode Decomposition and Extreme Learning Machines / S.M. Sulaiman // *Computers & Electrical Engineering*. – 2022. No 98. – pp. 107663.

144. Ning Xu., Yaoguo D., Yande G. Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China / Xu. Ning, D. Yaoguo, G. Yande // *Energy*. – 2017. – No 118. – pp. 473–480.

145. Chenyu Z., et.al. Using fractional discrete Verhulst model to forecast Fujian's electricity consumption in China / Z. Chenyu // *Energy*. – 2022. – No 255. – pp. 124484.

146. Nzoko T.D., et.al. Prediction of electrical energy consumption in Cameroon through econometric models / T.D. Nzoko // *Electric Power Systems Research*. – 2022. – No 210. – pp. 108102.

147. Palchevsky, E.V. Threat prediction in complex distributed systems using artificial neural network technology / E.V. Palchevsky, O.I. Khristodulo, S.V. Pavlov // *CEUR Workshop Proceedings*. – 2020. – Vol. 2763. – C. 289–284.

148. Wangwang Y., et.al. A combined deep learning load forecasting model of single household resident user considering multi-time scale electricity consumption behavior / Y. Wangwang // *Applied Energy*. – 2022. – No 307. – pp. 118197.

149. Ramos D., et.al. Using decision tree to select forecasting algorithms in distinct electricity consumption context of an office building / D. Ramos // *Energy Reports*. – 2022. – No 8–3. – pp. 417–422.

150. Xingcai Z., et.al. Panel semiparametric quantile regression neural network for electricity consumption forecasting / Z. Xingcai // *Ecological Informatics*. – 2022. – No 67. – pp. 101489.

151. Florencia L., et.al. User behaviour models to forecast electricity consumption of residential customers based on smart metering data / L. Florencia // Energy Reports. – 2022. – No 8. – pp. 3680–3691.

152. Knapinska A. Long-term prediction of multiple types of time-varying network traffic using chunk-based ensemble learning / A. Knapinska, P. Lechowicz, W. Wegier, K. Walkowiak // Applied Soft Computing. – 2022. – Vol. 130. – pp. 109694.

153. Постановление Правительства РФ от 04.05.2012 N 442 (ред. от 28.09.2023) «О функционировании розничных рынков электрической энергии, полном и (или) частичном ограничении режима потребления электрической энергии» – https://www.consultant.ru/document/cons_doc_LAW_130498/ (Дата обращения 19.12.2023)

154. Cron – <https://ru.wikipedia.org/wiki/Cron> (Дата обращения 28.06.2023).

155. FitzHugh R. Mathematical model of action potential and propagation in nerve / R. FitzHugh // Biological Engineering. – 1969. – pp. 1–85.

156. FitzHugh R. Impulses and physiological states in theoretical models of nerve membrane / R. FitzHugh // Biophys. J. – 1961. – no 1. – pp. 445–466.

157. McCulloch W.S., Pitts W. A logical calculus of the ideas immanent in nervous activity / W.S. McCulloch, W. Pitts // Bulletin of Mathematical Biophysics. – 1943. – No 5. – pp. 115–133.

158. Hodgkin A.L. A quantitative description of membrane current and its application to conduction and excitation in nerve / A. L. Hodgkin, A. F. Huxley // J. Physiol. – 1952. – No 117 (4). – pp. 500–544.

159. Пальчевский, Е.В. Разработка импульсной нейронной сети с возможностью скоростного обучения для нейтрализации DDoS-атак / Е.В. Пальчевский, О.И. Христовуло // Программные продукты и системы. – 2019. – Т. 32. № 4. – С. 613–627.

160. Mitchell T. Machine Learning. – 1997. – 432 p.

161. Fukushima K. Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements / K. Fukushima // *IEEE Transactions on Systems Science and Cybernetics*. – 1969. – No 5–4. – pp. 322–333.

162. Paninski L. Maximum likelihood estimation of cascade point-process neural encoding model / L. Paninski // *Network*. – 2004. – No 15. – pp. 243–262.

163. Palchevsky, E.V. A system based on an artificial neural network of the second generation for decision support in especially significant situations / E.V. Palchevsky, V.V. Antonov, R.R. Enikeev, T. Breikin // *Journal of Hydrology*. – 2023. – Vol. 616. – pp. 128844.

164. Palchevsky, E.V. Intelligent forecasting of electricity consumption in managing energy enterprises in order to carry out energy-saving measures / E.V. Palchevsky, V.V. Antonov, L.A. Rodionova, L.A. Kromina, A.R. Fakhrullina // *Mekhatronika, Avtomatizatsiya, Upravlenie*. – 2023. – Vol. 24. No 6. – pp. 307–316.

165. Palchevsky, E.V. The Concept of Formation of Intelligent Control Systems of Power Supply of Urban Networks / E.V. Palchevsky, V.V. Antonov, L.A. Rodionova, L.A. Kromina, A.R. Fakhrullina, L.I. Baimurzina, E.A. Rodionov // *Mekhatronika, Avtomatizatsiya, Upravlenie*. – 2023. – Vol. 24. No 4. – pp. 190–198.

166. Palchevsky, E.V. Intelligent system for monitoring and analyzing competencies in the learning process / E.V. Palchevsky, G.G. Kulikov, V.V. Antonov, L.E. Rodionova, A.R. Fakhrullina, L.A. Kromina // *Software & Systems*. – 2023. – Vol. 36. No 1. – pp. 5–13.

167. Пальчевский, Е.В. Разработка самообучающейся спайковой нейронной сети для проактивного реагирования на внешние информационные воздействия различной природы / Е.В. Пальчевский, В.В. Антонов, Л.Е. Родионова, Л.А. Кромина // *International Journal of Open Information Technologies (INJOIT)*. – 2022. – Т. 10. №7. – С. 1-12.

168. Palchevsky, E.V. Decision support system based on application of the second-generation neural network / E.V. Palchevsky, V.V. Antonov // *Programmная Ingeneria*. – 2022. – Vol. 13. No. 6. – С. 301–308.

169. Пальчевский, Е.В. Прогнозирование на основе искусственной нейронной сети второго поколения для поддержки принятия решений в особо значимых ситуациях / Е.В. Пальчевский, В.В. Антонов, Р.Р. Еникеев // Программные продукты и системы. – 2022. – Т. 35. № 3. – С. 488-503.

170. Palchevsky, E.V. Artificial Neural Network for Forecasting Electricity Consumption in Energy Enterprises / E.V. Palchevsky, V.V. Antonov, L.E. Rodionova, A.R. Fakhrullina, L.A. Kromina // Programmnaya Ingeneria. – 2023. – Vol. 14. No. 1. – pp. 34–41.

171. Palchevsky, E.V. Method for improving the accuracy of predictive values of time series based on the imputation of historical data / E.V. Palchevsky // Programmnaya Ingeneria. – 2023. – Vol. 14. No. 6. – pp. 301–306.

172. Пальчевский, Е.В. Анализ ретроспективных данных с применением технологий искусственного интеллекта для прогнозирования угроз в сложных распределенных системах / Е.В. Пальчевский, О.И. Христодуло, С.В. Павлов, А.В. Соколова // Вестник компьютерных и информационных технологий. – 2021. – №2. – С. 39–45.

173. Palchevsky, E.V. Intelligent data analysis for forecasting threats in complex distributed systems / E.V. Palchevsky, O.I. Khristodulo, S.V. Pavlov, A.M. Kalimgulov // CEUR Workshop Proceedings. – 2020. – Vol. 2744. – pp. 285–296.

174. Palchevsky, E.V. Threats complex distributed systems parrying based on their development prognostication / E.V. Palchevsky, O.I. Khristodulo, S.V. Pavlov // Advances in Social Science, Education and Humanities Research. – 2020. – Vol. 483. – pp. 191–194.

175. Пальчевский, Е.В. Прогнозирование угроз в сложных распределенных системах на основе интеллектуального анализа больших данных автоматизированных средств мониторинга / Е.В. Пальчевский, О.И. Христодуло, С.В. Павлов // Программные продукты и системы. – 2021. Т. 34. № 2. – С. 239-248.

176. Backpropagation – <https://builtin.com/machine-learning/backpropagation-neural-network> (Дата обращения 28.06.2023)

177. Интеллектуальная система прогнозирования «Elforecasting» – <https://elforecasting.com/predictEnergy?lang=ru> (Дата обращения 29.06.2023)

178. Большаков, А. А. Архитектура системы поддержки принятия решений при выборе конструктивных решений для автостереоскопических дисплеев / А. А. Большаков, А. В. Ключиков // Математические методы в технике и технологиях - ММТТ. – 2020. – Т. 12-1. – С. 35-38.

179. Степанов, В. В. Современные архитектуры интеллектуальных систем поддержки принятия решений / В. В. Степанов, К. М. Липин, И. Д. Коробейников // Электронный сетевой политематический журнал "Научные труды КубГТУ". – 2018. – № 3. – С. 324-334.

180. Мониторинг «Zabbix» – <https://www.zabbix.com/ru> (Дата обращения 29.06.2023)

ПРИЛОЖЕНИЕ А

(справочное)

Копии свидетельств о государственной регистрации программ для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2023662065

**«Программная библиотека глубокого обучения для
решения регрессионных задач»**

Правообладатель: *федеральное государственное бюджетное
образовательное учреждение высшего образования
«Уфимский университет науки и технологий» (RU)*

Авторы: *Пальчевский Евгений Владимирович (RU), Антонов
Вячеслав Викторович (RU)*

Заявка № 2023619608

Дата поступления 11 мая 2023 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 06 июня 2023 г.

Руководитель Федеральной службы
по интеллектуальной собственности



ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 4292640b738c1764b9f6183b73b4aa7
Идентификатор Юридический Сертификат
Действителен с 01.06.2023 по 02.08.2024

Ю.С. Зубов

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2017612872

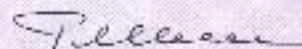
РАСЧЕТ КОЛИЧЕСТВА ВХОДЯЩИХ ПОТОКОВЫХ
СЕТЕВЫХ ПАКЕТОВПравообладатель: *Пальчевский Евгений Владимирович (RU)*Автор: *Пальчевский Евгений Владимирович (RU)*

Заявка № 2017610264

Дата поступления 10 января 2017 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 03 марта 2017 г.

Руководитель Федеральной службы
по интеллектуальной собственности
 Г.П. Новиков

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

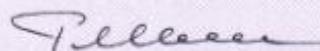
о государственной регистрации программы для ЭВМ

№ 2017613016

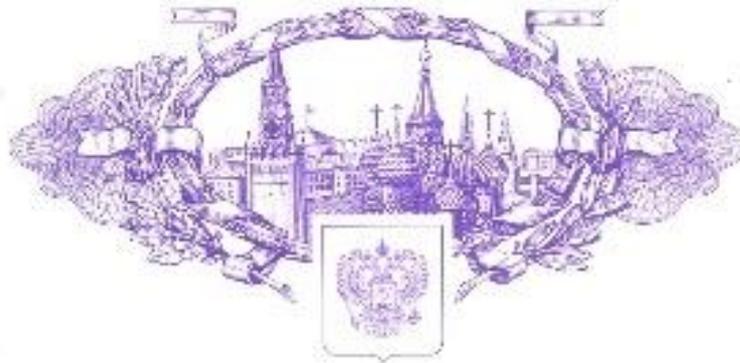
СИСТЕМА УПРАВЛЕНИЯ СЕРВЕРАМИ В
ВЕБ-ИНТЕРФЕЙСЕПравообладатель: *Пальчевский Евгений Владимирович (RU)*Автор: *Пальчевский Евгений Владимирович (RU)*

Заявка № 2017610137

Дата поступления 10 января 2017 г.

Дата государственной регистрации
в Реестре программ для ЭВМ 07 марта 2017 г.Руководитель Федеральной службы
по интеллектуальной собственности
 Г.П. Изhev


РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2017613065

РАСПАРАЛЛЕЛИВАНИЕ НАГРУЗКИ ФИЗИЧЕСКИХ И
ЛОГИЧЕСКИХ ЯДЕР

Правообладатель: *Пальчевский Евгений Владимирович (RU)*

Автор: *Пальчевский Евгений Владимирович (RU)*

Заявка № 2017610255

Дата поступления 10 января 2017 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 09 марта 2017 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Г.Н. Ильин

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2017613861

МОНИТОРИНГ СОСТОЯНИЯ ВЫЧИСЛИТЕЛЬНОГО ПРОЦЕССА

Правообладатель: *Пальчевский Евгений Владимирович (RU)*

Автор: *Пальчевский Евгений Владимирович (RU)*

Заявка № 2017610368

Дата поступления 09 января 2017 г.

Дата государственной регистрации

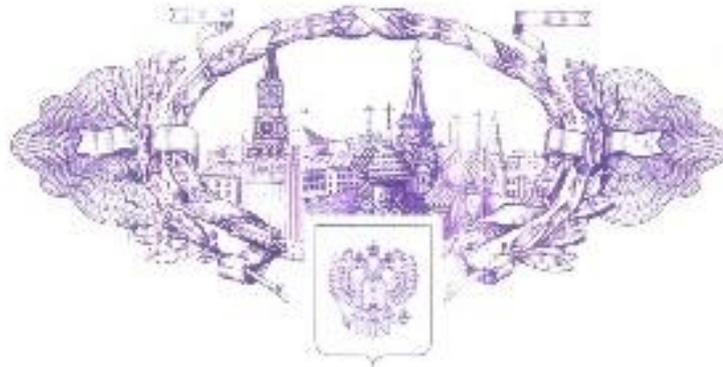
в Реестре программ для ЭВМ 03 апреля 2017 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Иванова Г.П. Иванова

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2017614111

РАЗРАБОТКА АЛГОРИТМА БЛОКИРОВКИ IP-АДРЕСА

Правообладатель: *Пальчевский Евгений Владимирович (RU)*Автор: *Пальчевский Евгений Владимирович (RU)*

Заявка № 2017610156

Дата поступления: 09 января 2017 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 06 апреля 2017 г.

Руководитель Федеральной службы
по интеллектуальной собственности
 Г.П. Иванов

ПРИЛОЖЕНИЕ Б
(справочное)

Копии актов о внедрении диссертационного исследования

УТВЕРЖДАЮ
Проректор по развитию
образования ФГБОУ ВО
«Уфимский университет науки и
технологий»
Рахманова Ю.В.
« 27 » 03 2023.



АКТ ВНЕДРЕНИЯ

О внедрении в учебный процесс университета результатов кандидатской диссертации аспиранта кафедры «Автоматизированных систем управления» Пальчевского Евгения Владимировича по теме «Методы нейросетевой обработки больших темпоральных данных для информационной поддержки принятия управленческих решений (на примере электроэнергетики)»

Мы, нижеподписавшиеся, и.о. декана факультета информатики и робототехники, к.т.н. Ковтуненко А.С., заместитель заведующего кафедрой «Автоматизированных систем управления» к.т.н. Суворова В.А., составили настоящий акт о том, что полученные аспирантом кафедры «Автоматизированных систем управления» Пальчевским Е.В. основные результаты кандидатской диссертации (метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений) внедрены в учебный процесс университета и применяются в учебном процессе для решения задач моделирования и управления в дисциплинах «Математическое моделирование сложных систем» и «Специальные информационно-аналитические системы», где требуется:

1. Строить нечеткие системы из наборов цифровых данных в рамках лабораторных и практических работ.
2. Реализовывать практические инструменты по моделированию различных процессов и прогнозированию необходимых параметров в рамках курсовых работ.

И.о. декана факультета
информатики и робототехники
к.т.н.



А.С. Ковтуненко

Зам. заведующего кафедрой
«Автоматизированных систем управления»
к.т.н.



В.А. Суворова



ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ
«ГАЗПРОМНЕФТЬ – ЦИФРОВЫЕ РЕШЕНИЯ»
(ООО «ГАЗПРОМНЕФТЬ – ЦР»)

Адрес для корреспонденции:
ул. Киевская, д. 5, корп. 4, Санкт-Петербург, 196004
ОКПО 056603189, ОГРН 1087746449312,
ИНН 7728054530, КПП 781001001
Тел.: +7 812 448-24-01
e-mail: ds-info@gazprom-neft.ru
ds.gazprom-neft.ru

УТВЕРЖДАЮ

Руководитель направления управления
технологического развития
«ООО Газпромнефть – Цифровые решения»
А.С. Дьячков



«02 марта 2023 г.»

АКТ

о внедрении результатов диссертационной работы
Пальчевского Евгения Владимировича

на тему: «Методы нейросетевой обработки больших темпоральных данных для
информационной поддержки принятия управленческих решений (на примере
электроэнергетики)»

Председатель комиссии: Иван Сергеевич Вологин, Главный аналитик интеграций,
Члены комиссии: Айдар Радикович Акназаров, Руководитель направления
ресурсного обеспечения,
Ильнур Ильдарович Исянчурин, Заместитель
руководителя центра - Руководитель практики поставки
данных

Настоящим актом подтверждаем, что в «Практике интеграционных решений и экспертизы» ООО «Газпромнефть – Цифровые решения» внедрены и используются предложенные в диссертационном исследовании Пальчевского Е.В. метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений.

В основу программного решения по мониторингу оборудования обеспечивающего обработку потоков данных из разных источников с получением прогнозных показателей в процессе обеспечения интеллектуального информационного сопровождения, положен метод нейросетевой обработки больших данных и прогнозирования временных рядов, базирующийся на предложенных модели рекуррентной нейронной сети и ее модифицированном методе обучения.

Отмечается низкое потребление вычислительных ресурсов – не более 6% центрального процессора и 4% накопителей, что дает возможность обработки больших объемов информации без дополнительного резервирования вычислительных мощностей и автоматического генерирования правил фильтрации потока по реализации упреждающего реагирования на аномалии в сетевом потоке данных и очистку данных для формирования датасета.

Главный аналитик интеграций

И.С. Вологин

Руководитель направления
ресурсного обеспечения

А.Р. Акназаров

Главный инженер данных

И.И. Исянчурин

УТВЕРЖДАЮ

Заместитель главного инженера

Приуфимской ТЭЦ

Начальник площадки Павловская ГЭС

Иванов А.Н.

«16» марта 2023 г.

**АКТ**

о внедрении результатов диссертационной работы
Пальчевского Евгения Владимировича

на тему: «Методы нейросетевой обработки больших темпоральных данных для информационной поддержки принятия управленческих решений (на примере электроэнергетики)»

Комиссия в составе:

председатель: Заместитель главного инженера ПУТЭЦ-
начальник площадки Павловская ГЭС Иванов А.Н.

члены комиссии: начальник ЭМЦ Кириллов Б.Р.
начальник участка АСУ Данилов С.В.

Мы, нижеподписавшиеся, составили настоящий акт о том, что следующие результаты диссертационной работы использовались при подготовке технических требований к разработке и внедрению автоматизированной системы предиктивной аналитики и диагностики, основанной на комплексной оценке и анализе состояния и режимных параметров гидроагрегатов для площадки Павловская ГЭС Приуфимской ТЭЦ филиала ООО «БГК»:

1. Метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений.
2. Метод и модель обработки информации сетевого трафика с применением импульсной нейронной сети для подготовки исходного датасета данных с целью прогнозирования значений потребления электроэнергии.

Результаты диссертационного исследования Пальчевского Е.В. внедрены при разработке научно-обоснованной методологии построения предиктивной аналитики и диагностики, основанной на комплексной оценке и анализе вибрационного состояния и режимных параметров гидроагрегата в виде структуры цифрового двойника ГА, объединенных в программный комплекс (систему) предиктивной аналитики и диагностики на основе:

- формального описания объектов, их взаимодействия и процессов управления данными;
- разработки формализованных правил объединения данных для мониторинга оборудования из разных источников (баз данных АСУТП и из файлов разных форматов);
- создание классификаторов дефектов и формализованных рекомендаций (правил) по их устранению или переключения от одного вида ремонта к другому;
- создание прогнозных моделей на основе применений нейронной сети и оценки остаточного ресурса на основе динамики развития ситуации потенциальных инцидентов (отказов контролируемых узлов агрегатов) – прогноз состояния оборудования;
- интеграции прогнозных показателей с принятыми решениями для обеспечения интеллектуального информационного сопровождения.

Начальник ЭМЦ



Кириллов Б.Р.

Начальник участка АСУ



Данилов С.В.

УТВЕРЖДАЮ

ООО «Башкирэнерго»
ПО «Кумертауские электрические
сети»

Заместитель директора по
электросетевым услугам,
начальник отдела транспорта
электрической энергии

К.Л. Богатырев

2023 г.

**АКТ****о внедрении результатов диссертационной работы**

Пальчевского Евгения Владимировича

на тему: «Методы нейросетевой обработки больших темпоральных данных для информационной поддержки принятия управленческих решений (на примере электроэнергетики)»

Председатель комиссии: Д.В. Марфин, заместитель начальника отдела транспорта электрической энергии

Члены комиссии: В.Е. Гальцев, начальник отдела присоединения.

Мы, нижеподписавшиеся, составили настоящий акт о том, что следующие результаты диссертационной работы приняты к внедрению в производственное отделение «Кумертауские электрические сети» ООО «Башкирэнерго»:

1. Метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети для информационной поддержки принятия управленческих решений.
2. Структурная схема системы управления процессом поддержки принятия управленческих решений в электроэнергетике, включающая информационную поддержку на основе предложенного математического аппарата комплекса моделей и методов.

Метод и модель прогнозирования многомерных временных рядов с применением рекуррентной нейронной сети использованы в программном обеспечении регламентного мониторинга оборудования и системы датчиков.

Методика управления процессом информационного сопровождения математического аппарата комплекса моделей и методов разработана с учетом принципа построения контура управления, включающего структурную схему системы управления процессом поддержки принятия управленческих решений.

Эффективность обусловлена снижением рисков для ряда задач в контуре управления по несвоевременному принятию решений за счет повышения быстродействия и точности (в среднем до 92%) получения прогнозных показателей без необходимости предварительной перегрузки данных в специализированное хранилище с дополнительной загрузкой вычислительных мощностей.

Заместитель начальника отдела
транспорта электрической энергии

 Д.В. Марфин

Начальник отдела присоединения

 В.Е. Гальцов