

Архитектура доверенной системы искусственного интеллекта в задачах прогнозирования временных рядов в электроэнергетике

П. В. Матренин

Аннотация—В условиях роста электропотребления и развития распределенной генерации в энергетических системах возрастает потребность в надежных и объяснимых системах прогнозирования. В статье предложена новая архитектура доверенной системы искусственного интеллекта для прогнозирования временных рядов в электроэнергетике. В отличие от существующих решений, ориентированных преимущественно на точность моделей, предложенная архитектура направлена на реализацию доверенности системного уровня на основе онтологического подхода, а также обеспечения информационной безопасности и генерации объяснений на естественном языке. Разработанная архитектура включает модули контроля целостности входных данных и моделей, детектирования дрейфа и автоматической адаптации моделей, а также модуль объяснений, использующий языковую модель, обученную на корпусе объяснений, соответствующих векторам значений признаков Шепли. Онтологическая модель, интегрированная с общей информационной моделью, обеспечивает согласованность компонентов системы. Обоснованы требования к доверенным системам искусственного интеллекта в задачах прогнозирования временных рядов в электроэнергетике, включая безопасность, объяснимость, защиту от деградации моделей и совместимость с информационными системами объектов энергетики. Представленные решения могут быть применены для прогнозирования нагрузки, генерации электрических станций на базе возобновляемых источников энергии и иных задач принятия решений при планировании и управлении на объектах энергетики. Результаты обосновывают возможность практической реализации систем, обеспечивающих доверие как отдельных пользователей, так и индустрии в целом на основе принципов онтологического моделирования.

Ключевые слова—архитектура информационной системы, доверенная система искусственного интеллекта, прогнозирование временных рядов в электроэнергетике, объяснимый искусственный интеллект.

I. ВВЕДЕНИЕ

Системы прогнозирования временных рядов, использующие модели искусственного интеллекта,

становятся важным инструментом обеспечения эффективного планирования режимов работы электроэнергетических систем.

Рост доли распределенной генерации, возобновляемых источников энергии (ВИЭ), рыночных механизмов обеспечения баланса генерации и потребления приводят к существенному увеличению объемов данных и уровню неопределенности при решении задач планирования работы объектов энергетики. Качество планирования напрямую зависит от точности и достоверности краткосрочных и среднесрочных прогнозов, как отмечается во многих работах, таких как [1-6]. Существующие модели на основе машинного обучения демонстрируют высокие показатели точности в лабораторных стационарных условиях, однако их слабая интерпретируемость и недостаточная защищенность от деградации моделей из-за дрейфа данных не способствуют доверию пользователей, таких как системные операторы, энергосбытовые и генерирующие компании, крупные потребители электроэнергии [6, 7].

Понятие «доверенный искусственный интеллект» может быть определено как совокупность технологических, организационных и этических свойств интеллектуальной системы, гарантирующих ее безопасность, прозрачность и соответствие нормативным требованиям на всем жизненном цикле. В российской практике требования к доверенным системам указаны в ГОСТ Р 59276–2020 «Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения» и ГОСТ Р 59898-2021 «Оценка качества систем искусственного интеллекта. Общие положения».

Существующие решения, как правило, фокусируются на одном из трех аспектов: повышение точности путем все большего усложнения моделей [3, 6]; объяснимость формируемых результатов [7-9]; обновление моделей [10-12].

Как правило, архитектуры информационных систем прогнозирования в задачах электроэнергетики включают слой или модуль предобработки данных, модуль интеллектуального анализа и модуль постобработки [13, 14, 15]. Могут быть выделены различные внешние информационные системы, такие как метеопровайдеры, ERP- и SCADA-системы [15, 16].

Статья получена 4 мая 2025.

П. В. Матренин, канд. тех. наук, в.н.с., p.v.matrenin@urfu.ru
Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», Екатеринбург.

Можно отметить статьи [17, 18], в которых предобработка служит не только для исправления ошибок, заполнения пропусков, но и для контроля целостности, доступности и достоверности входных данных.

В статье [19] предложена фреймворк на основе принципов MLOps не только для разработки моделей, но для их надежной эксплуатации.

Перспективным является подход, основанный на использовании онтологических моделей в архитектуре систем прогнозирования выработки ВИЭ, описанный в исследовании [20]. В этой работе предлагается промежуточный слой, который использует онтологию в качестве модели семантической интеграции гетерогенных данных от различных источников. Онтологический подход как средство для построения информационных систем в электроэнергетике также представлен в работе [21].

Можно выделить следующие основные направления исследований в области архитектур интеллектуальных систем прогнозирования в электроэнергетике:

- добавление модулей объяснения результатов моделей; как правило, используются аддитивный алгоритм объяснения Шепли (SHAP) и алгоритм локальных интерпретируемых моделей (LIME);
- реализация принципов MLOps для автоматического обновления моделей, контроля качества работы систем;
- применение семантических моделей для интеграции разнородных данных.

Целью данного исследования является разработка архитектуры доверенной системы искусственного интеллекта для прогнозирования временных рядов в электроэнергетике на основе обеспечения доверенности на трех уровнях: семантическом (использование онтологических моделей); технологическом (модульная интеграция подсистемы защиты и адаптации); функциональном (формирование объяснений, механизмы контроля дрейфа данных и автоматического обновления моделей).

II. СИСТЕМАТИЗАЦИЯ ТРЕБОВАНИЙ К ИНТЕЛЛЕКТУАЛЬНЫМ СИСТЕМАМ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ В ЭЛЕКТРОЭНЕРГЕТИКЕ

Несмотря на разнообразие, существующие архитектуры интеллектуальных систем прогнозирования временных рядов в электроэнергетике обладают следующими недостатками:

- Слабая семантическая управляемость системы. Отсутствие единой онтологии приводит к дублированию сущностей и ошибкам при интеграции с ERP/SCADA-системами.
- Ограниченная объяснимость моделей ИИ. Объяснение реализуется на уровне таких методов как SHAP и LIME, формирующих лишь вектор значимостей признаков и соответствующие визуализации, которые являются понятными скорее разработчикам моделей, чем конечным пользователям.
- Отсутствие отдельного модуля детектирования дрейфа данных и деградации моделей. Разнообразие видов дрейфа требует разработки стратегии

обновления моделей [12, 22, 23].

- Обеспечение контроля целостности и доступности только входных данных, без защиты систем от угроз подмены обучающей выборки или модели.

Опираясь на нормативные документы и анализ практик операторов сетей, сформулированы следующие сквозные требования к архитектуре.

1. Обеспечение защиты информационной безопасности на уровне входных данных. Контроль входных данных (доступности и целостности), детектирование аномалий и дрейфа данных.
2. Контроль целостности модели ИИ (защита от подмены модели или ее обучаемых параметров).
3. Объяснимость результатов с помощью генерации человеко-читаемых объяснений в текстовом и визуальном форматах, раскрывающих факторы, повлиявшие на прогноз.
4. Адаптивность к дрейфу данных. Система должна включать онлайн-детектор статистических сдвигов и поддерживать автоматическое обновление модели.
5. Возможность интеграции с внешними информационными системами путем использования общих информационных моделей, таких как CIM/IEC 61970/61968.

III. ПРЕДЛОЖЕННАЯ АРХИТЕКТУРА СИСТЕМЫ

В качестве основы при проектировании архитектуры использован онтологический подход, поскольку он позволяет корректно учитывать особенности предметной области и в то же время онтологические модели достаточно формализованы, чтобы использовать их математически при реализации системы. Кроме того, онтологический подход основан на принципах системного анализа и является одним из его важных направлений.

Целостность интеллектуальной системы прогнозирования временных рядов в электроэнергетике невозможна без единых принципов описания информации, передающейся между компонентами архитектуры. В данной работе основу такой целостности составляет внутренняя онтологическая модель, которая может быть интегрирована с отраслевым стандартом CIM/IEC 61970 и адаптирована под специфику производственных и прочих факторов, влияющих на генерацию и потребление электроэнергии.

На Рис. 1 показана концепция разработанной архитектуры доверенной интеллектуальной системы прогнозирования временных рядов в задачах электроэнергетики на базе онтологического подхода и принципов системного анализа. Ее отличиями от представленных в литературе архитектур является наличие модулей контроля целостности моделей искусственного интеллекта, детектирования дрейфа данных, адаптации моделей, а также объяснения результатов на естественном языке с использованием языковой модели (Рис. 2). Все указанные дополнения направлены на достижение единой цели – повышение доверенности системы (Табл. 1).

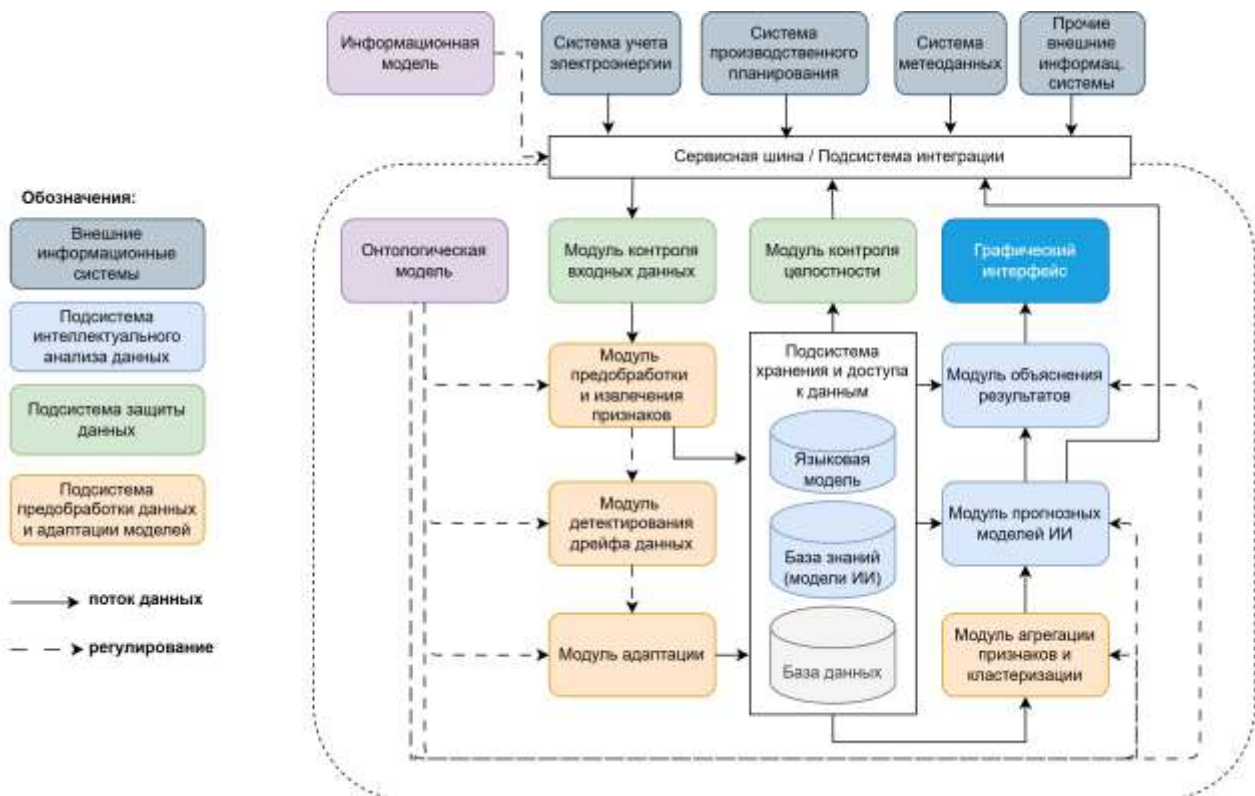


Рис. 1: Концепция архитектуры доверенной интеллектуальной системы прогнозирования временных рядов в задачах электроэнергетики на базе онтологического подхода

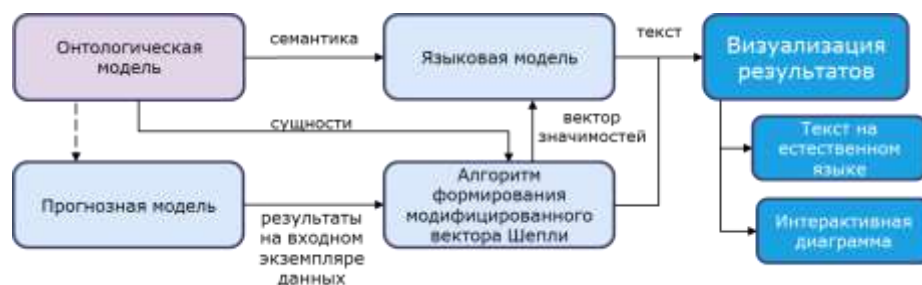


Рис. 2: Концептуальная схема предложенного метода формирования объяснений прогнозов

Модуль предобработки и извлечения признаков является стандартным, аналогичным описанным в работах [13-17] и отличается от них только использованием онтологической модели, задающей состав извлекаемых признаков.

Модули детектирования дрейфа данных и адаптации могут быть реализованы согласно описаниям в работах [10-12, 22]. Подробный обзор методов детектирования дрейфа и рекомендаций по их выбору приведен к работе [22]. В случае детектирования дрейфа выполняется асинхронное повторное обучение или дообучения модели и ее тестирование. Это выполняется на копии модели, поэтому не нарушает процесса функционирования системы, даже при ее работе в режиме приближенном к реальному времени. Затем выполняется горизонтальное обновление, новая модель разворачивается параллельно действующей,

сравнивается по ключевым метрикам, и после успешной верификации происходит замена экземпляров моделей.

Модуль контроля входных данных выполняет проверки на целостность и доступность данных, детектирование пропусков и ошибок. Данные разделены на базу данных, базу знаний (модели искусственного интеллекта и значения их обучаемых параметров), и языковую модель для формирования объяснений на естественном языке.

В данной архитектуре подсистема хранения и доступа к данным – это логически объединенная, но физически распределенная система, включающая: базу данных для хранения результатов измерений, метаданных, погодных факторов и производственных данных. Предлагается использовать реляционную базу данных.

База знаний представляет из себя систему сериализованных моделей ИИ и их обученных

параметров в формате ONNX для нейросетевых моделей и Joblib для прочих моделей, таких как ансамбли деревьев решений, а также краткие описания моделей, составленные разработчиками в формате MD.

Табл. 1: Влияние модулей архитектуры на обеспечение доверенности интеллектуальной системы

Компонент архитектуры	Обеспечиваемое свойство доверия
Модуль контроля входных данных	Защита целостности и доступности входной информации; фильтрация аномалий и недостоверных значений
Модуль контроля целостности модели	Защита от подмены модели ИИ и ее параметров; контроль версий
Онтологическая модель	Семантическая согласованность, устранение дублирования понятий, автоматическая генерация API
Модуль извлечения и агрегации признаков	Адаптивность к структуре данных и новизне объектов; онтологически контролируемое формирование признаков
Детектор дрейфа данных	Обнаружение дрейфа данных
Модуль адаптации моделей	Автоматическое обновление ИИ-моделей для адаптации к дрейфу данных
Прогнозный модуль	Обеспечение высокой точности и достоверности прогнозов
Языковая модель	Объяснимость: генерация понятных объяснений на естественном языке
Сервисная шина	Возможность масштабируемой, безопасной и стандартизированной интеграции с внешними информационными системами

В качестве языковой модели предлагается использование компактных языковых моделей, таких как mT5-small [23]. Данная модель представляет собой компактный вариант семейства трансформерных моделей с архитектурой encoder-decoder, реализованной на основе механизма внимания. Подсистема хранения и доступа к данным отвечает за ее хранение в формате ONNX. Подробнее использование языковой модели описано ниже.

Модуль агрегации и кластеризации признаков позволяет строить не единую прогнозную модель для всех возможных входных данных, а несколько моделей для различных кластеров, что повышает итоговую точность прогнозов и снижает риски переобучения моделей [24, 25].

Модуль прогнозных моделей непосредственно строит прогнозы, которые затем объясняются модулем объяснений. Как было отмечено ранее, доверенная система искусственного интеллекта должна не только формировать прогнозы с требуемой точностью, но и

объяснять свои решения пользователю на понятном ему языке. Результаты объяснения позволяют специалисту проанализировать влияние факторов на результат прогнозирования и понять логику построения прогноза, а система в этом случае становится для пользователя уже не черным ящиком, а понятным инструментом и помощником.

Для формирования объяснений может быть использован алгоритм SHAP. Аддитивное объяснение Шепли представляет собой алгоритм определения вклада каждого признака в результат модели для каждого отдельного входного экземпляра [8, 9].

Но базовый алгоритм Шепли дает значимости признаков, которые на самом деле сложно понимать пользователям. Поэтому предложен метод генерации объяснения, понятного пользователю, на естественном языке. Онтология в этом случае позволяет сделать объяснение не просто на естественном языке (русском, английском), но на языке предметной области.

Сервисная шина или подсистема интеграции также направлена на повышение качества интеграции. Используется тот же принцип замены взаимодействия point-to-point на общение через разделяемую систему обмена данными, что и при использовании CIM-моделей в электроэнергетике или при создании цифровых подстанций на основе стандарта IEC 61850.

IV. ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ В СИСТЕМЕ

Онтологическая модель обеспечивает целостность системы и ее согласованность с предметной областью, поскольку все сущности, метаданные и атрибуты в рамках системы оказываются объединены единым семантическим пространством.

В архитектуре используется внутренняя онтологическая модель, формализованная в формате OWL (Web Ontology Language), которая описывает сущности предметной области и определяет структуру и семантическую спецификацию признаков. Онтология включает такие классы как ElectricLoad (электропотребление), AirTemperature (температура воздуха), OperationalSchedule (производственный план), HeatingSystemLoad (нагрузка на систему обогрева) и другие, описывающие объекты и их свойства, влияющие на процессы генерации и потребления электроэнергии.

Для каждого класса в онтологии заданы:

- атрибуты, такие как среднечасовая температура воздуха для AirTemperature, значение электропотребления для ElectricLoad и другие, при этом у атрибутов заданы не только значения, но и единицы измерения и типы данных;
- связи между сущностями, например, ElectricLoad зависит от HeatingSystemLoad и OperationalSchedule, а HeatingSystemLoad зависит от AirTemperature).

На этапе предобработки система использует онтологию как декларативное описание схемы признаков. Механизм семантического сопоставления позволяет автоматически извлекать признаки из входных данных, на основе сопоставления входных данных с сущностями онтологии согласно приведенному ниже алгоритму.

Общая информационная модель, при ее наличии на объекте, для которого разрабатывается система прогнозирования, служит для упрощения интеграции, ускорения процесса внедрения и снижения рисков ошибок благодаря единому семантическому аппарату и общим сущностям при программной реализации системы.

Алгоритм 1. Семантическое сопоставление признаков

```

Вход:
    O – онтология с описанием сущностей и признаков
    D – описание структуры входных данных
    M – таблица соответствий

НАЧАЛО

f = []

Для каждого класса C в O:
    Для каждого атрибута A в C:
        Если есть C.A. в M:
            s = M[C.A]  # s – имя поля в источнике данных

            Если есть s в D:
                f = f ∪ {
                    "f_name": C.name + "." + A.name,
                    "data_field": s,
                    "type": A.type,
                    "unit": A.unit}

Вернуть f

КОНЕЦ

```

Механизм формирования объяснений в предложенной архитектуре строится на согласованном взаимодействии трёх компонентов: онтологической модели, алгоритма объяснения SHAP и языковой модели mT5-small.

Онтология описывает предметную область в виде семантической структуры: сущности (классы), их атрибуты, связи между ними. Она служит основой для интерпретации результатов прогноза с точки зрения смыслов, привычных специалистам в предметной области.

Алгоритм SHAP используется для оценки вклада каждого признака в конкретный прогноз. Он обеспечивает количественную значимость признаков, но не даёт естественно-языкового объяснения.

семантический анализ связей между сущностями в онтологии позволяет выделить цепочку влияния признаков на целевую переменную и интерпретировать численные значения SHAP-вектора в терминах предметной области. Например, если SHAP указывает на высокий вклад признака «прогнозная температура воздуха» в прогноз электропотребления, то с помощью онтологической модели будет выявлена кратчайшая цепочка связей: «погода влияет на загрузку системы отопления, которая влияет на электропотребления». Языковая модель, дообученная на корпусе объяснений, преобразует семантическую структуру объяснения в текстовое объяснение, понятное человеку. При этом отбираются признаки, значимость которых по модулю

не ниже определенного порога (3–7 %).

Поэтому модель необходимо дообучить на датасете, состоящем из элементов вида $\{F, V, U, R, E\}$, где F – наименования признаков согласно онтологии; V – значения признаков, U – значимости признаков, полученные с помощью алгоритма SHAP; R – цепочки зависимостей целевой переменной от признаков (извлекается из онтологической модели путем поиска кратчайшего пути от признака к целевой переменной, например, AirTemperature – HeatingSystemLoad – ElectricLoad); E – составленные экспертами тексты объяснений для обучения модели.

Описанный подход позволяет реализовать следующие свойства:

1. Семантическая непротиворечивость. Каждый объект системы связан с онтологическим классом, что исключает дублирование, неоднозначность и противоречивые трактовки сущностей. Все данные, модели и отчеты опираются на единое семантическое пространство.
2. Верифицируемость. Связь компонентов архитектуры с онтологией позволяет отслеживать происхождение и логику каждого результата прогноза. Например, прогнозное значение нагрузки связано с определенным классом ElectricLoad, основано на наборах признаков WeatherFactor и OperationalSchedule, каждый из которых верифицируется через онтологический слой.
3. Автоматическая генерация классов данных и API. На основе онтологической модели автоматически формируются структуры классов и их атрибутов в коде, а также OpenAPI-спецификации для REST/gRPC-интерфейсов. Автоматическая генерация кода снижает трудоемкость при масштабировании системы и обеспечивает согласованность всех слоев: данных, логики и интерфейсов.
4. Семантический вывод и объяснимость. В системе реализован модуль, использующий семантический анализ онтологической модели для генерации объяснений на естественном языке.

V. ЗАКЛЮЧЕНИЕ

Предложенная в статье архитектура доверенной системы искусственного интеллекта ориентирована на отраслевые особенности электроэнергетики, где принятие решений требует высокой точности прогнозов, строгой верифицируемости источников информации, устойчивости к изменяющимся условиям, а также объяснения рекомендуемых решений. Электроэнергетика как отрасль характеризуется не только критичностью принимаемых решений, но и высокой степенью регламентированности, необходимостью обеспечения кибербезопасности и интеграции с уже существующими информационными системами, такими как SCADA, ERP, АСУ ТП [26, 27].

Онтологический подход позволяет при разработке информационных систем сформировать единое семантическое пространство, объединяющее данные, модели, интерфейсы и объяснения. Онтологическая модель служит не просто справочной структурой, а активным ядром архитектуры, сквозным способом координации всех уровней системы — от данных и прогнозных моделей до логики объяснения и взаимодействия с внешними системами, что соответствует принципам системного анализа, где каждый элемент оценивается в контексте целого, а целостность системы обеспечивается не компоновкой функций, а единством модели мира.

Предложенная архитектура может быть применена в задачах краткосрочного прогнозирования нагрузки, генерации на объектах с ВИЭ, предиктивной диагностики и управления режимами. В качестве дальнейших направлений исследований предполагается: разработка метода многокритериальной формализованной количественной оценки уровня доверенности систем искусственного интеллекта; а также развитие и обоснование метода формирования объяснений на естественном языке.

Статья демонстрирует, как онтологическая модель может быть практически интегрирована в архитектуру интеллектуальной системы и использоваться не только для структурирования данных, но и для повышения объяснимости и доверия к результатам прогнозов.

БЛАГОДАРНОСТИ

Работа выполнена в рамках государственного задания при финансовой поддержке Министерства науки и высшего образования Российской Федерации (тема № FEUZ-2025-0005 Разработка моделей и методов объяснимого искусственного интеллекта для повышения надежности и безопасности внедрения распределенных интеллектуальных систем на объектах электроэнергетики).

БИБЛИОГРАФИЯ

- [1] Моргоев И. Д., Дзгоев А. Э., Клюев Р. В., А. Д. Моргоева. Прогнозирование потребления электроэнергии предприятиями народнохозяйственного комплекса в условиях неполноты информации // *Известия Кабардино-Балкарского научного центра РАН*, 2022, № 3 (107), С. 9–20.
- [2] Макоклюев Б. И. Методы и программные средства планирования электропотребления // *Энергия единой сети*, 2022, № 2, С. 58–69.
- [3] Ahmad N., Ghadi Y., Adnan M., Ali M. Load Forecasting Techniques for Power System: Research Challenges and Survey // *IEEE Access*, 2022, vol. 10, pp. 71054-71090.
- [4] Блохин А. В., Грицай А. С., Горшенин А. Ю. Исследование факторов, влияющих на потребление электроэнергии коммерческим предприятием // *Математические структуры и моделирование*, 2022, № 3(63), С. 39–47.
- [5] Глазирин А. С., Боловин Е. В., Архипова О. В. [и др.]. Адаптивное краткосрочное прогнозирование потребления электроэнергии автономными энергосистемами малых северных поселений на основе методов ретроспективного регрессионного анализа // *Известия Томского политехнического университета. Инжиниринг георесурсов*, 2023, Т. 334, № 4, С. 231–248.
- [6] Сергеев Н. Н., Матренин П. В. Обзор международного опыта в прогнозировании генерации возобновляемых источников энергии с помощью методов машинного обучения // *iPolytech Journal*, 2023, Т. 27, № 2. С. 354-369.

- [7] Горшенин А. Ю. Формирование выборки исходных данных для машинного обучения модели краткосрочного прогнозирования электропотребления // *Автоматизация в промышленности*, 2023, № 10, С. 37–41.
- [8] Baur L., Ditschuneit K., Schambach M. [et al]. Explainability and Interpretability in Electric Load Forecasting Using Machine Learning Techniques – A Review // *Energy and AI*, 2024, vol. 16, art. 100358.
- [9] Matrenin P. V., Stepanova A. I. Enhancing the interpretability of electricity consumption forecasting models for mining enterprises using SHapley Additive exPlanations // *Journal of Mining Institute*, 2025, vol. 271, pp. 154–167.
- [10] Rebei A., Amayri M., Bouguila N. Affinity-Driven Transfer Learning for Load Forecasting // *Sensors*, 2024, vol. 24 (17), art. 5802.
- [11] Bin Kamilin M. H., Yamaguchi S., Bin Ahmadon M. A. Radian Scaling and Its Application to Enhance Electricity Load Forecasting in Smart Cities Against Concept Drift // *Smart Cities*, 2024, vol. 7, pp. 3412–3436.
- [12] Матренин П. В., Хамитов Р. Н. “Алгоритм формирования стратегии автоматического обновления моделей искусственного интеллекта в задачах прогнозирования в электроэнергетике, *Инженерный вестник Дона*, 2025.
- [13] Habibi M. R., Golestan S., Wu Y., Guerrero J. M., Vasquez J. C. Electrical load forecasting in power systems based on quantum computing using time series-based quantum artificial intelligence // *Scientific Reports*, 2025, vol 15, art. 7429.
- [14] Miraki A., Parviainen P., Arghandeh R. Electricity demand forecasting at distribution and household levels using explainable causal graph neural network // *Energy and AI*, 2024, vol. 16, art. 100368.
- [15] Горшенин А. Ю. Разработка и верификация программного обеспечения системы краткосрочного прогнозирования электропотребления // *Известия Тульского государственного университета. Технические науки*, 2024, № 7, С. 253–259.
- [16] Antonesi G., Cioara T., Todorean L., Anghel I., De Mulder C. A Machine Learning Pipeline to Forecast the Electricity and Heat Consumption in a City District // *Buildings*, 2023, vol 13, art. 1407.
- [17] Xu H., Hu F., Liang X. A framework for electricity load forecasting based on attention mechanism time series depthwise separable convolutional neural network // *Energy*, 2024, vol. 299, art. 131258.
- [18] Khalyasmaa A., Matrenin P. Initial Data Corruption Impact on Machine Learning Models' Performance in Energy Consumption Forecast // *2021 Ural-Siberian Smart Energy Conference (USSEC)*, Novosibirsk, Russian Federation, 2021.
- [19] Pelekis S., Karakolis E., Pountridis T. [et al]. DeepTSF: Codeless machine learning operations for time series forecasting. URL <https://arxiv.org/abs/2308.00709> (дата обращения: 23.04.2025).
- [20] Wicaksono H., Trat M., Bashyal A. [et al]. Artificial-intelligence-enabled dynamic demand response system for maximizing the use of renewable electricity in production processes // *The International Journal of Advanced Manufacturing Technology*, 2025, vol. 138, pp. 247–271.
- [21] Массель Л. В., Массель А. Г. Семантическое моделирование при построении цифровых двойников энергетических объектов и систем // *Онтология проектирования*, 2023, № 1(13), С. 44-52.
- [22] Намиот Д.Е., Ильюшин Е.А. Мониторинг сдвига данных в моделях машинного обучения // *International Journal of Open Information Technologies*, 2022, Т. 10, № 12. С. 84–93.
- [23] Xue L., Constant N., Roberts A. [et al.]. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer // *2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 483–498.
- [24] Matrenin P. V., Khalyasmaa A. I., Gamaley V. V. [et al]. Improving of the Generation Accuracy Forecasting of Photovoltaic Plants Based on k-Means and k-Nearest Neighbors Algorithms // *Energetika. Proceedings of CIS Higher Education Institutions and Power Engineering Associations*, 2023, vol. 66(4), pp. 305–321.
- [25] Zhang Z., Wang C., Peng X. [et al]. Solar Radiation Intensity Probabilistic Forecasting Based on K-Means Time Series Clustering and Gaussian Process Regression // *IEEE Access*, 2021, vol. 9, pp. 89079–89092.
- [26] Khalyasmaa A. I., Matrenin P. V. New software architecture for visualization, transformation, and analysis of equipment images // *International Journal of Open Information Technologies*, 2025, Vol. 13 (5), pp. 34–40.
- [27] Arrano-Vargas F., Konstantinou G. Modular Design and Real-Time Simulators Toward Power System Digital Twins Implementation // *IEEE Transactions on Industrial Informatics*, 2023, vol. 19(1), pp. 52–61.

Architecture of a Trusted Artificial Intelligence System for Time Series Forecasting in Power Industry

P.V. Matrenin

Abstract— As energy consumption increases and distributed generation develops, the need for reliable and explainable forecasting systems becomes critical. The new article presents the architecture of a trusted artificial intelligence system for time series forecasting in power industry. Unlike existing solutions that primarily focus on model accuracy, the proposed approach implements system-level trust based on ontological models, information security, and natural language explanation generation. The developed architecture includes modules for input data integrity control, model verification, data drift detection, automatic model adaptation, and a dedicated explanation module. The latter employs a language model trained on a corpus of human-authored explanations aligned with Shapley values. The ontological model, integrated with a common information model, ensures semantic consistency across system components. The paper substantiates the key requirements for trusted AI systems in time series forecasting for the power sector, including security, explainability, protection against model degradation, and compatibility with enterprise information systems. The proposed solutions can be applied to load forecasting, renewable generation prediction, and decision support in power system planning and operation. The results demonstrate the practical feasibility of implementing AI systems that foster trust—both at the level of individual users and across the industry—by relying on the principles of ontological modeling.

Keywords—information system architecture, trusted artificial intelligence system, time series forecasting in power industry, explainable artificial intelligence.

REFERENCES

- [1] I. D. Morgoyev, A. E. Dzgoyev, R. V. Klyuyev, and A. D. Morgoyeva, "Prognozirovaniye potrebleniya elektroenergii predpriyatiyami narodnokhozyaystvennogo kompleksa v usloviyakh nepolnoty informatsii [Forecasting electricity consumption by enterprises of the national economic complex in conditions of incomplete information]," *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN [Bulletin of the Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences]*, vol. 3 (107), pp. 9–20, 2022.
- [2] B. I. Makoklyuyev, "Metody i programmnyye sredstva planirovaniya elektropotrebleniya [Methods and software for planning electricity consumption]," *Energiya yedinoi seti [Energy of the unified network]*, vol. 2, pp. 58–69, 2022.
- [3] N. Ahmad, Y. Ghadi, M. Adnan and M. Ali, "Load Forecasting Techniques for Power System: Research Challenges and Survey," *IEEE Access*, vol. 10, pp. 71054–71090, 2022.
- [4] A. V. Blokhin, A. S. Gritsay, and A. Y. Gorshenin, "Issledovaniye faktorov, vliyayushchikh na potrebleniye elektroenergii kommercheskim predpriyatiyem [Study of factors influencing electricity consumption by a commercial enterprise // Mathematical structures and modeling.]," *Matematicheskiye struktury i modelirovaniye [Mathematical structures and modeling]*, vol. 3 (63), pp. 39–47, 2022.
- [5] A. S. Glazyrin, E. V. Bolovin, O. V. Arkhipova [et al] "Adaptive short-term forecasting of electricity consumption by autonomous power systems of small northern settlements based on retrospective regression analysis methods," *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, vol. 334, no. 4, pp. 231–248, 2023.
- [6] P. V. Matrenin and N. N. Sergeyev, "Obzor mezhdunarodnogo opyta v prognozirovanii generatsii vozobnovlyayemykh istochnikov energii s pomoshch'yu metodov mashinnogo obucheniya [Review of international experience in forecasting the generation of renewable energy sources using machine learning methods]," *iPolytech Journal*, vol. 27(2), pp. 354–369, 2023.
- [7] A. Y. Gorshenin, "Formirovaniye vyborki iskhodnykh dannykh dlya mashinnogo obucheniya modeli kratkosrochnogo prognozirovaniya elektropotrebleniya [Formation of a sample of initial data for machine learning of a short-term forecasting model of electricity consumption]," *Avtomatizatsiya v promyshlennosti [Automation in industry]*, vol. 10, pp. 37–41, 2023.
- [8] L. Baur, K. Ditschuneit, and M. Schambach [et al], "Explainability and Interpretability in Electric Load Forecasting Using Machine Learning Techniques – A Review," *Energy and AI*, vol. 16, art. 100358, 2024.
- [9] P. V. Matrenin and A. I. Stepanova, "Enhancing the interpretability of electricity consumption forecasting models for mining enterprises using SHapley Additive exPlanations," *Journal of Mining Institute*, vol. 271, pp. 154–167, 2025.
- [10] A. Rebei, M. Amayri, N. Bouguila, "Affinity-Driven Transfer Learning for Load Forecasting," *Sensors*, vol. 24 (17), art. 5802, 2024.
- [11] M. H. Bin Kamilin, S. Yamaguchi, and M. A. Bin Ahmadon, "Radian Scaling and Its Application to Enhance Electricity Load Forecasting in Smart Cities Against Concept Drift," *Smart Cities*, vol. 7, pp. 3412–3436, 2024.
- [12] P. V. Matrenin and R. N. Khamitov, "Algoritm formirovaniya strategii avtomaticheskogo obnoveniya modelyey iskusstvennogo intellekta v zadachakh prognozirovaniya v elektroenergetike [Algorithm for forming a strategy for automatic updating of artificial intelligence models in forecasting problems in the electric power industry]," *Inzhenernyy vestnik Dona [Engineering Bulletin of the Don]*, 2025.
- [13] M. R. Habibi, S. Golestan, Y. Wu, J. M. Guerrero, and J. C. Vasquez, "Electrical load forecasting in power systems based on quantum computing using time series-based quantum artificial intelligence," *Scientific Reports*, vol. 15, art. 7429 (2025).
- [14] A. Miraki, P. Parviainen, and R. Arghandeh, "Electricity demand forecasting at distribution and household levels using explainable causal graph neural network," *Energy and AI*, vol. 16, art. 100368, 2024.
- [15] A. Y. Gorshenin, "Razrabotka i verifikatsiya programmno obespicheniya sistemy kratkosrochnogo prognozirovaniya elektropotrebleniya [Development and verification of software for a short-term forecasting system for electricity consumption]," *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskiye nauki [Bulletin of Tula State University. Technical sciences]*, vol. 7, pp. 253–259, 2024.
- [16] G. Antonesi, T. Cioara, L. Todorean, I. Anghel, and C. De Mulder, "A Machine Learning Pipeline to Forecast the Electricity and Heat Consumption in a City District," *Buildings*, vol. 13, art. 1407, 2023.
- [17] H. Xu, F. Hu, and X. Liang, "A framework for electricity load forecasting based on attention mechanism time series depthwise separable convolutional neural network," *Energy*, vol. 299, art. 131258, 2024.

- [18] A. Khalyasmaa and P. Matrenin, "Initial Data Corruption Impact on Machine Learning Models' Performance in Energy Consumption Forecast," *2021 Ural-Siberian Smart Energy Conference (USSEC)*, Novosibirsk, Russian Federation, 2021.
- [19] S. Pelekis, E. Karakolis, T. Pountridis [et al], "DeepTSF: Codeless machine learning operations for time series forecasting," URL <https://arxiv.org/abs/2308.00709> (access date: 23.04.2025).
- [20] H. Wicaksono, M. Trat, A. Bashyal [et al], "Artificial-intelligence-enabled dynamic demand response system for maximizing the use of renewable electricity in production processes," *The International Journal of Advanced Manufacturing Technology*, vol. 138, pp. 247–271, 2025.
- [21] L. V. Massel and A. G. Massel, "Semanticheskoye modelirovaniye pri postroyenii tsifrovyykh dvoynikov energeticheskikh ob'yektov i sistem [Semantic modeling in constructing digital twins of energy objects and systems]," *Ontologiya proyektirovaniya [Design Ontology]*, vol. 1(13), pp. 44-52, 2023.
- [22] D. E. Namiot and E. A. Ilyushin, "Monitoring sdviga dannykh v modelyakh mashinnogo obucheniya [Monitoring data shift in machine learning models]," *International Journal of Open Information Technologies*, vol. 10(12), pp. 84-93, 2022.
- [23] Xue L., Constant N., Roberts A. [et al.], "mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer," *2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 483–498.
- [24] P. V. Matrenin, A. I. Khalyasmaa, V. V. Gamaley [et al], "Improving of the Generation Accuracy Forecasting of Photovoltaic Plants Based on k-Means and k-Nearest Neighbors Algorithms," *Energetika. Proceedings of CIS Higher Education Institutions and Power Engineering Associations*, vol. 66(4), pp. 305–321, 2023.
- [25] Z. Zhang, C. Wang, X. Peng [et al], "Solar Radiation Intensity Probabilistic Forecasting Based on K-Means Time Series Clustering and Gaussian Process Regression," *IEEE Access*, vol. 9, pp. 89079–89092, 2021.
- [26] A. I. Khalyasmaa and P. V. Matrenin, "New software architecture for visualization, transformation, and analysis of equipment images," *International Journal of Open Information Technologies*, 2025, Vol. 13 (5), pp. 34–40.
- [27] F. Arrano-Vargas and G. Konstantinou, "Modular Design and Real-Time Simulators Toward Power System Digital Twins Implementation," *IEEE Transactions on Industrial Informatics*, 2023, vol. 19(1), pp. 52–61.