

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА 24.2.437.10,
СОЗДАННОГО НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО
АВТОНОМНОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО
ОБРАЗОВАНИЯ «ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ (НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ)» МИНИСТЕРСТВА НАУКИ И ВЫСШЕГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ, ПО ДИССЕРТАЦИИ
НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ КАНДИДАТА НАУК

аттестационное дело № _____

решение диссертационного совета от 20 мая 2026 г. № 1

О присуждении Юртину Алексею Артемьевичу, гражданину Российской Федерации, ученой степени кандидата физико-математических наук.

Диссертация «Нейросетевые методы восстановления потоковых данных» по специальности 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей принята к защите 11 марта 2026 г. (протокол заседания № 3) диссертационным советом 24.2.437.10, созданным на базе федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)», Министерство науки и высшего образования Российской Федерации, 454080, г. Челябинск, пр. Ленина, 76, приказ от 12 октября 2022 г. № 1170/нк.

Соискатель Юртин Алексей Артемьевич, 11 августа 1998 года рождения, в 2020 г. и 2022 г. окончил федеральное государственное автономное образовательное учреждение высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)» по направлениям подготовки 09.03.01 «Информатика и вычислительная техника» и 02.04.02 «Фундаментальная информатика и информационные технологии». В 2025 г. окончил очную аспирантуру

федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)» по научной специальности 2.3.5 Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

В настоящее время соискатель работает программистом в Лаборатории больших данных и машинного обучения, а также на условиях внутреннего совместительства преподавателем на кафедре системного программирования в федеральном государственном автономном образовательном учреждении высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)» Министерства науки и высшего образования Российской Федерации.

Диссертация выполнена на кафедре системного программирования федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)» Министерства науки и высшего образования Российской Федерации.

Научный руководитель – доктор физико-математических наук, доцент Цымблер Михаил Леонидович, заместитель директора Научно-образовательного центра «Искусственный интеллект и квантовые технологии», федеральное государственное автономное образовательное учреждение высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)».

Официальные оппоненты:

- Громов Василий Александрович, доктор физико-математических наук, профессор, заместитель руководителя департамента анализа данных и искусственного интеллекта, федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет «Высшая школа экономики» (г. Москва),

- Воеводин Вадим Владимирович, кандидат физико-математических наук, заведующий Лабораторией анализа суперкомпьютерных систем и приложений Научно-исследовательского вычислительного центра, федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В. Ломоносова» (г. Москва),

дали положительные отзывы на диссертацию.

Ведущая организация – федеральное государственное автономное образовательное учреждение высшего образования «Казанский (Приволжский) федеральный университет» (г. Казань) в своем положительном отзыве, подписанном Елизаровым Александром Михайловичем, доктором физико-математических наук, профессором, профессором кафедры цифровой аналитики и технологий искусственного интеллекта Института информационных технологий и интеллектуальных систем, и утвержденном Таюрским Д.А., доктором физико-математических наук, профессором, первым проректором - проректором по научной деятельности, указала, что диссертационная работа А.А. Юртина представляет собой самостоятельную и законченную научно-квалификационную работу, в которой разработаны новые нейросетевые методы восстановления многомерных временных рядов, обладающие значимой практической ценностью для решения задач обработки потоковых данных. По уровню решаемых задач, содержанию и полученным результатам диссертационная работа соответствует требованиям, установленным Положением о порядке присуждения ученых степеней, в том числе п. 9, а ее автор заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Соискатель имеет 7 опубликованных работ, в том числе по теме диссертации опубликовано 5 работ, из них в рецензируемых научных изданиях

опубликовано 5 работ. Все 5 работ опубликованы в российских научных журналах Перечня ВАК (К-1, К-2).

Публикации полностью соответствуют теме диссертационного исследования и раскрывают ее основные положения. В диссертации отсутствуют недостоверные сведения об опубликованных соискателем ученой степени работах, в которых изложены основные научные результаты диссертации. Все научные результаты, выносимые на защиту, получены соискателем лично.

Наиболее значимые научные работы соискателя по теме диссертации:

1. Yurtin A.A., Zymbler M.L. SANNI: Online Imputation of Missing Values in Multivariate Time Series Based on Deep Learning and Behavioral Patterns // Lobachevskii Journal of Mathematics. 2024. Vol. 45, no. 11. P. 5948–5966. DOI: 10.1134/S1995080224606854. WOS: 001446922600003. (Перечень ВАК К1, Scopus Q2) (15 с./19 с.). Соискателю принадлежат разделы 1-5: обзор работ, формальная постановка задачи, описание метода восстановления временных рядов, описание результатов экспериментов, заключение). М.Л. Цымблеру принадлежит введение.
2. Юртин А.А. Метод прогнозирования ошибки времени обучения нейросетевых моделей восстановления многомерных временных рядов // Проблемы информатики. 2025. № 3. С. 78–95. DOI: 10.24412/2073-0667-2025-3-72-95. (Перечень ВАК К2).
3. Юртин А.А. Об одной функции потерь для обучения нейросетевых моделей восстановления временных рядов // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. Т. 13, № 4. С. 53–73. DOI: 10.14529/cmse240404. (Перечень ВАК К2).
4. Юртин А.А. Восстановление многомерных временных рядов на основе выявления поведенческих шаблонов и применения автоэнкодеров // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. Т. 13, № 2. С. 39–55. DOI: 10.14529/cmse240203. (Перечень ВАК К2).

5. Цымблер М.Л., Юртин А.А. Восстановление пропущенных значений временного ряда на основе совместного применения аналитических алгоритмов и нейронных сетей // Вычислительные методы и программирование. 2023. Т. 24, № 3. С. 243–259. DOI: 10.26089/NumMet.v24r318. (Перечень ВАК К1) (14 с./18 с.). Соискателю принадлежат разделы 1-5: обзор работ, формальная постановка задачи, описание метода восстановления временных рядов, описание результатов экспериментов, заключение). М.Л. Цымблеру принадлежит введение.

На диссертацию и автореферат поступили положительные отзывы:

1. Ведущей организации, ФГАОУ ВО «Казанский (Приволжский) федеральный университет» (г. Казань). Замечания:
 - 1) Новый метод SANNI восстановления временного ряда работает в режиме онлайн и при длительных сериях пропусков использует предыдущие восстановленные значения. Это может приводить к накоплению ошибки. Необходимо пояснить, почему предложенный метод является устойчивым к длительным сериям пропусков и каков механизм контроля надежности восстановленных значений.
 - 2) Новая функция потерь MPDE позволяет обучать нейросетевые модели с учетом поведенческого сходства подпоследовательностей временного ряда. Почему же в вычислительных экспериментах качество нейросетевых моделей, обученных с использованием MPDE, оценивается с помощью другой, традиционной метрики RMSE, которая не использовалась в процессе обучения.
 - 3) Некоторые части текста дублируют друг друга, особенно когда идет перечень того, что будет обсуждено или рассмотрено ниже. Все такие информативные части (где сказано, что и где будет в диссертации далее) есть во введении, они достаточно подробно составлены и, как нам представляется, не требуют повторений. Например, вводный текст в раздел на стр. 32 без ущерба можно было перенести на стр. 22

или вообще исключить. То же относится к информативному материалу на стр. 68.

- 4) В диссертации (глава 3) есть Лемма 1 и несколько утверждений. Эта лемма представляется тривиальной и доказательства не требует (речь идет о простом выводе простого числового неравенства). Последующие утверждения как математические объекты также являются простыми и получаются несложными выкладками. И лемма, и названные утверждения несколько не повышают математические уровень и значимость диссертации и без ущерба могли бы быть просто переведены в технические выкладки.
 - 5) В тексте достаточно часто встречаются банальные или, наоборот, совсем непонятные фразы, которые требуют многократного прочтения для понимания и, скорее всего, переписывания. Например, «... поскольку квадрат любого действительного числа неотрицателен» (стр. 87); «уменьшаемое второго множителя правой части формулы ...», «в соответствии с цепным правилом ...» (стр. 92); «для наивного (прямого) вычисления вклада» (стр. 93); «частная производная MPDE по переменной вычисляется как среднее значение производных расстояний между окнами» (стр. 94); «как сумма разностей между точками восстановленного и истинного временного ряда, деленных на евклидово расстояние соответствующих окон ...» (стр. 95); «входными данными метода являются архитектура целевой нейросетевой модели и вектор параметров ее обучения» (стр. 101). Текст диссертации, в целом написанный грамотно и хорошим языком, выиграл бы, если бы таких фраз было меньше.
2. Официального оппонента, доктора физико-математических наук, профессора, заместителя руководителя департамента анализа данных и искусственного интеллекта ФГАОУ ВО «Национальный исследовательский университет «Высшая школа экономики» (г. Москва) Громова Василия Александровича. Замечания:

- 1) Работа, безусловно, выиграла бы от понимания того, насколько зависит эффективность предложенных методов от степени хаотичности рассматриваемых рядов (значения старшего показателя Ляпунова, положение пары значений энтропия-сложность ряда на соответствующей плоскости и др.).
 - 2) Насколько возможен перенос рассмотренных в работе подходов на ряды других предметных областей?
 - 3) Каковы минимальные объёмы выборок, на которых методы демонстрируют свою эффективность?
3. Официального оппонента, кандидата физико-математических наук, заведующего Лабораторией анализа суперкомпьютерных систем и приложений Научно-исследовательского вычислительного центра ФГБОУ ВО «Московский государственный университет имени М.В. Ломоносова» (г. Москва) Воеводина Вадима Владимировича.

Замечания:

- 1) В разделах 2 и 3, посвященных описанию архитектуры разрабатываемых методов (а именно, SANNI, SAETI и tsGAP), не дано пояснения, почему в каждом случае была выбрана именно такая архитектура нейросети (например, почему автоэнкодер был выбран для SAETI, но не для SANNI). Также не указано, пробовались ли другие виды нейросетей и/или другие варианты их архитектур.
- 2) В разделе 4.3.1 не пояснено, почему метод SANNI проверялся на четырех сценариях формирования пропусков, а SAETI – только на одном. Аналогичный вопрос относительно функции потерь MDPE, которая проверялась только на сценарии Blackout (раздел 4.4.1).
- 3) Для функции потерь MDPE необходимо вручную подбирать значения параметров α и β , что несколько затрудняет работу с ней. Стоило привести в работе рекомендации по подбору этих значений либо реализовать возможности для автоматизации этого процесса.

- 4) В тексте диссертации присутствуют небольшие неточности. Например, в Заключение сказано, что «Вычислительные эксперименты, проведенные на реальных и синтетических данных, продемонстрировали превосходство метода tsGAP над передовыми аналогами в среднем на 38.8%», однако ранее нигде в тексте диссертации это значение не встречается. Также в тексте присутствуют пунктуационные ошибки.
4. Доктора физико-математических наук, профессора, заведующего лабораторией математической химии, ведущего научного сотрудника Института нефтехимии и катализа Уфимского федерального исследовательского центра РАН (г. Уфа) Губайдуллина Ирека Марсовича. Замечания отсутствуют.
5. Кандидата технических наук, доцента, заведующего кафедрой высокопроизводительных вычислений и системного программирования ФГАОУ ВО «Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского» (г. Нижний Новгород) Меерова Иосифа Борисовича. Замечания:
- 1) Характеризуя разработанный параллельный алгоритм, было бы полезно указать параметры его масштабируемости.
 - 2) Также имеет смысл привести данные о вычислительных ресурсах, необходимых при решении прикладных задач, в зависимости от их трудоемкости.
6. Доктора технических наук, доцента, ведущего научного сотрудника Научно-исследовательского центра супер-ЭВМ и нейрокомпьютеров (г. Таганрог) Никитиной Аллы Валерьевны. Замечания отсутствуют.
7. Кандидата технических наук, доцента, старшего научного сотрудника исследовательского центра в сфере искусственного интеллекта «Сильный искусственный интеллект в промышленности» ФГАОУ ВО «Национальный исследовательский университет ИТМО» (г. Санкт-Петербург) Семёнова Александра Анатольевича. Замечания:

- 1) В работе восстановление неполных временных рядов осуществляется только с использованием нейросетевого подхода. Хотелось бы увидеть сравнение полученных результатов с результатами, которые можно получить за счет применения других известных ML-моделей (например, для рассмотренных задач явно бы подошли деревья решений или случайные леса). По крайней мере, не помешало бы аргументированное обоснование, почему именно нейросети являются наиболее подходящим выбором в данном контексте.
- 2) Утверждение 4 на стр. 15 автореферата относительно «вычислительной сложности функции MDPE» нуждается в уточнении: что понимается под элементарными операциями, суммарное число которых характеризует сложность предложенного алгоритма?

Выбор официальных оппонентов и ведущей организации обосновывается широкой известностью их научных достижений в областях интеллектуального анализа временных рядов, машинного обучения, прогнозирования и восстановления потоковых данных, направленных на применение методов искусственного интеллекта для мониторинга и анализа сложных систем, включая суперкомпьютерные комплексы и физические процессы. В ведущей организации действует Институт искусственного интеллекта, робототехники и системной инженерии.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

- **разработан** комплекс нейросетевых методов восстановления пропущенных значений в потоковых данных, представленных в форме многомерных временных рядов, обеспечивающих повышение точности восстановления по сравнению с существующими аналогами на данных, содержащих повторяющиеся поведенческие шаблоны;
- **предложена** функция потерь MPDE для обучения нейросетевых моделей восстановления временных рядов, учитывающая

поведенческое сходство между подпоследовательностями и обеспечивающая повышение качества восстановления по сравнению с традиционными функциями потерь;

- **разработан** параллельный алгоритм вычисления функции MPDE, позволяющий использовать ее в качестве функции потерь при обучении нейросетевых моделей восстановления потоковых данных в современных фреймворках глубокого обучения;
- **разработан** нейросетевой метод tsGAR для прогнозирования ошибки и времени обучения нейросетевых моделей восстановления потоковых данных в форме многомерных временных рядов, отличающийся совместным кодированием графового представления целевой нейросетевой модели и гиперпараметров обучения с последующей агрегацией признаков, превосходящий передовые аналоги по точности прогнозирования ошибки.

Теоретическая значимость исследования обоснована тем, что:

- **доказаны** утверждения о свойствах функции MPDE, позволяющие использовать ее в качестве функции потерь при обучении нейросетевых моделей восстановления потоковых данных;
- **применительно к проблематике диссертации результативно использованы** методы глубокого обучения, интеллектуального анализа данных, теория временных рядов, концепция поведенческих шаблонов и методы параллельного программирования;
- **изложена** идея использования поведенческих шаблонов многомерного временного ряда совместно с нейросетевыми моделями для повышения точности восстановления пропусков;
- **раскрыты** ограничения существующих нейросетевых методов восстановления пропущенных значений, не учитывающих поведенческие шаблоны временных рядов; несоответствие между традиционными точечными функциями потерь и требованием сохранения локальной формы подпоследовательностей; ограничения

традиционных функций потерь нейросетевых моделей восстановления потоковых данных, не обеспечивающих сохранение поведенческого сходства подпоследовательностей;

- **изучены** связь между долей пропущенных значений и точностью восстановления при использовании поведенческих шаблонов; зависимость производительности параллельного алгоритма вычисления функции MPDE от длины подпоследовательности.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

- **разработаны и опубликованы** в сети Интернет в репозиториях GitVerse для свободного использования исходные тексты, реализующие предложенные в диссертации нейросетевые модели, методы и алгоритмы;
- **определена** эффективность разработанных нейросетевых моделей, методов и функции потерь в сравнении с известными аналогами путем проведения вычислительных экспериментов на реальных и синтетических временных рядах из различных предметных областей в соответствии с общепринятыми стандартами и сценариями формирования пропусков;
- **созданы** новые нейросетевые модели, методы и алгоритмы восстановления пропущенных значений в потоковых данных, представленных в форме многомерных временных рядов, которые могут использоваться в качестве инструментальных средств разработки на различных этапах жизненного цикла моделей машинного обучения;
- **представлены** результаты экспериментальных исследований, подтверждающие способность метода SANNI обрабатывать данные в режиме реального времени с задержкой, удовлетворяющей требованиям систем промышленной автоматизации и автоматизации зданий, и демонстрирующие практическую применимость разработанных решений в интеллектуальных системах обработки потоковых данных.

Оценка достоверности результатов исследования выявила, что:

- **теория** построена на известных методах машинного обучения, глубокого обучения, теории временных рядов и параллельных вычислений;
- **идея базируется** на критическом анализе научных публикаций по теме исследования в высокорейтинговых отечественных и зарубежных журналах;
- **результаты получены** на платформе современных высокопроизводительных вычислительных систем с использованием современных фреймворков глубокого обучения;
- **использованы** общепринятые сценарии тестирования методов восстановления потоковых данных, стандартные метрики качества, открытые наборы данных и современные средства разработки нейросетевых моделей, обеспечивающие воспроизводимость экспериментов;
- **установлено** количественное и качественное соответствие результатов проведенных вычислительных экспериментов исходным данным на основе общепринятых метрик оценки точности восстановления.

Личный вклад соискателя состоит в том, что:

- **соискателем единолично разработаны** все полученные результаты, выносимые на защиту, в том числе: нейросетевые модели и методы восстановления потоковых данных в режимах онлайн и офлайн, новая функция потерь и параллельный алгоритм ее вычисления, нейросетевой метод прогнозирования ошибки и времени обучения нейросетевых моделей восстановления потоковых данных;
- в публикациях, выполненных в соавторстве с научным руководителем, научному руководителю принадлежит постановка задачи, **соискателю принадлежат все полученные результаты.**

В ходе защиты диссертации критические замечания высказаны не были, были заданы следующие вопросы:

1. Рассматривалась ли возможность хранения извлеченных шаблонов с использованием современных векторных баз данных (например, PostgreSQL) и поиска не по точному совпадению, а по близости результата запроса?
2. За счет чего предложенная функция потерь MPDE обеспечивает более высокую точность восстановления по сравнению с Dynamic Time Warping (DTW) и другими метриками, если DTW хорошо улавливает форму временных рядов?
3. В методах SANNI и SAETI используется нейросетевая модель «Распознаватель», которая обучается на выборках, полученных с помощью алгоритма поиска шаблонов. Возникает риск, что одни шаблоны встречаются часто, а другие – редко (дисбаланс классов). Может ли выборка быть несбалансированной, и если да, то как это учитывается?

Соискатель Юртин А.А. ответил на задаваемые ему в ходе заседания вопросы и привел собственную аргументацию:

1. В рамках текущего исследования шаблоны используются как часть нейросетевой модели. Как одно из направлений будущих исследований рассматривается возможность организации хранилища шаблонов с поиском по сходству с помощью нейросетевых моделей, что позволит избежать привязки к количеству шаблонов и добавлять их динамически.
2. MPDE вычисляет два вида расстояний: евклидово между точками окна и расстояние между нормализованными окнами. При нормализации окна ошибка в одной точке меняет среднее и стандартное отклонение всего окна, из-за чего изменяются нормализованные значения соседних точек, увеличивая ошибку для всего окна.
3. Да, дисбаланс возможен, и он учитывается. В случае если выборка не сбалансирована, используется два способа: применяется искусственная балансировка путем аугментации редких шаблонов и используется взвешенная функция потерь (веса обратно пропорциональны частоте класса).

На заседании 20 мая 2026 г. диссертационный совет принял решение за разработку новых нейросетевых методов восстановления потоковых данных, представленных в форме временных рядов, которые имеют существенное значение для повышения эффективности интеллектуального анализа данных и систем автоматизированного машинного обучения, присудить Юртину А.А. ученую степень кандидата физико-математических наук.

При проведении тайного голосования диссертационный совет в количестве 14 человек, из них 8 докторов наук по научной специальности рассматриваемой диссертации, участвовавших в заседании, из 17 человек, входящих в состав совета, дополнительно введены на разовую защиту 0 человек, проголосовали: за 14, против 0.

Председатель
диссертационного совета



Соколинский Леонид Борисович

Ученый секретарь
диссертационного совета



Краева Яна Александровна

20 мая 2026 г.

