

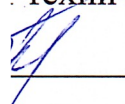
УТВЕРЖДАЮ

Проректор по научной работе
и международной деятельности



ФГБОУ ВО «Юго-Западный
государственный университет»

технических наук, доцент

 Е.Г. Пахомова

сентябрь 2024 года

ОТЗЫВ

ведущей организации о диссертации

Кожемякина Леонида Валерьевича

*«Методы и алгоритмы предобработки ограниченных данных
и оценки качества неполносвязных нейронных сетей
при их построении на основе корней принятия решений»,*

представленной на соискание ученой степени кандидата технических наук
по специальности 2.3.1. Системный анализ, управление и обработка
информации, статистика

1. Актуальность темы диссертации

Актуальность разработки новых методов и алгоритмов предварительной обработки ограниченных данных не вызывает сомнений, поскольку качество аппроксимации напрямую связано как с объёмом обучающих данных, так и с выбором необходимой структуры нейронной сети, которые в свою очередь являются одним из эффективных инструментов машинного обучения.

Однако для достижения эффективности применения нейронных сетей требуется достаточная и качественная выборка исходных данных, что не всегда является тривиальной и возможной задачей для ее формирования. Поэтому, немаловажным является то, что обработка исходных данных занимает ключевую позицию в общем алгоритме обработки информации и анализа данных, в том числе с применением аппарата искусственных нейронных сетей.

Вышесказанное подтверждает актуальность темы исследования.

2. Структура и содержание диссертации

Диссертационная работа по содержанию соответствует требованиям к научно-квалификационным работам на соискание ученой степени кандидата наук. Диссертационная работа имеет внутреннее единство и логически структурирована, включает введение, четыре главы, заключение, список сокращений и условных обозначений, список литературы и приложения.

Во введении описывается актуальность темы исследования, степень разработанности, объект, предмет, цель и задачи исследования, положения, выносимые на защиту, обладающие научной новизной, теоретическая и практическая значимость результатов диссертационной работы, методы исследования, степень достоверности теоретических результатов и выводов, апробация результатов, количество опубликованных научных работ.

В первой главе представлен обзор методов предварительной обработки информации в условиях ограниченных наборов данных. Делается акцент на двух методах, которые в совокупности позволяют перейти к структуре неполносвязной нейронной сети, основанной на корне принятия решений, выбор параметров и последовательность свёртки которого определяются с помощью выявления информационной значимости этих параметров в результате системно-когнитивного анализа.

Во второй главе выполнен обзор методов нейросетевых технологий, включая простые нейронные сети с одним скрытым слоем и сложные нейронные сети с каскадной и свёрточной структурой. Предложена блок-схема, предлагаемых правил и методов предварительной обработки данных и выбора архитектуры нейронной сети для обучения нейронной сети, основанной на корне принятия решений (нейронная сеть класса DRB NN) в случаях ограниченного набора данных.

В третьей главе рассмотрены информационные системы обработки информации, которые соответствуют предложенным правилам и методам предварительной обработки данных и выбора архитектуры нейронной сети класса DRB NN. Показан разработанный программный модуль для проектирования и

обучения нейронных сетей класса DRB NN, который использует матрицы свёртки для определения структуры неполносвязной ИНС. Один из функционалов предложенного программного модуля позволяет применять метод перекрёстной проверки (кросс-валидации) для оценки обобщающей способности прогностических моделей.

Четвертую главу следует считать апробацией разработанных методов и алгоритмов предварительной обработки ограниченных данных на примере анализа медицинских данных. Последовательность применения предлагаемых методов и алгоритмов позволила определить наиболее значимые факторы в исходном наборе данных, что определило вид корня принятия решений, на основе которого была получена структура неполносвязной нейронной сети. Оценка качества неполносвязной нейронной сети проводилась с использованием перекрёстной проверки. Качество полученной модели сравнивалось с качеством нейронных сетей с одним скрытым слоем и с двумя скрытыми слоями. Оказалось, что качество неполносвязных нейронных сетей в среднем превосходит качество полносвязных нейронных сетей, сопоставимых по числу синаптических коэффициентов, идентифицируемых в процессе обучения. Показано сравнение с другим инструментом анализа данных – деревья решений, как одним из видов машинного обучения, что также подтверждает, что неполносвязная нейронная сеть класса DRB NN обладает приемлемым качеством.

В заключении констатируется факт достижения цели исследования, приведены основные выводы по диссертационной работе в разрезе поставленных задач и полученных научных результатов, а также приводятся рекомендации и перспективы дальнейшей разработки темы.

Список использованной литературы включает в себя 187 источников и отражает современное состояние дел в исследуемой проблемной области. На все источники даны ссылки по тексту. По итогам анализа приведенных источников можно утверждать, что проведен хороший анализ литературы, о чем также свидетельствует содержание раздела «Степень разработанности проблемы».

В приложениях показано использование результатов диссертации при выполнении фундаментальных исследований по Технологическому проекту «Инновационные химические, медицинские и фармацевтические технологии» по подпроекту «Новые материалы и технологии для медицины». Также представлены копии свидетельств о регистрации трёх программ для ЭВМ, копия патента на изобретение.

3. Научная новизна и достоверность полученных результатов

В рамках исследования Кожемякин Л.В. получил следующие результаты, характеризующиеся научной новизной:

1. Предложены новые правила и методы обработки ограниченных наборов данных, согласно которым области определения исходных данных разбиваются на интервалы, образуя набор дискретных данных, далее идентифицируются корни принятия решений и на их основе строятся неполносвязные нейронные сети, которые обучаются на исходных данных. Такие правила и методы обработки ограниченных данных позволяют сократить поиск требуемой структуры неполносвязной нейронной сети.

2. Разработано алгоритмическое обеспечение информационной системы проектирования и обучения искусственных нейронных сетей на основе корней принятия решений. Информационная система позволяет реализовать многократные вычислительные эксперименты для анализа устойчивости результатов обучения и тестирование неполносвязных нейронных сетей.

3. Предложен метод оценки качества обучения нейронной сети в условиях ограниченных данных. Эксперименты показали, что качество неполносвязных сетей в среднем выше, чем у полносвязных сетей с аналогичным числом синаптических коэффициентов, идентифицируемых в процессе обучения.

Научная новизна подтверждается отличием разработанных методов проектирования и обучения нейронных сетей на основе корней принятия решений. Основные положения и выводы, полученные в диссертационной работе, являются обоснованными и аргументированными. Достоверность результатов подтверждена в ходе многочисленных экспериментов.

4. Теоретическая значимость результатов, полученных в диссертации

Теоретическая значимость обусловлена эффективностью разработанных соискателем новых методов и алгоритмов оценки качества неполносвязных нейронных сетей на основе корней принятия решений, в условиях ограниченного набора исходных данных.

5. Апробация и практическая значимость результатов, полученных в диссертации

Судя по списку работ, опубликованных соискателем по теме диссертации, апробация результатов диссертации проводилась на различных международных и всероссийских конференциях. В рамках диссертационной работы было создано программное обеспечение, которое используется как составной компонент виртуальной среды интеллектуального анализа данных «Data to Decisions» (D2D.Platform), разработанное ООО «Пермский центр поддержки принятия решений» в рамках НИОКР «Разработка и тестирование прототипа многопользовательской виртуальной среды интеллектуального анализа данных». Также программное обеспечение использовалось при выполнении фундаментальных исследований в рамках соглашений в 2022 и 2023 гг. с АНО «Пермский научно-образовательный центр мирового уровня «Рациональное недропользование» по проекту «Новые материалы и технологии для медицины».

Достоверность результатов и выводов подтверждается многократными вычислительными экспериментами, выполненными с применением различных методов обработки данных, используя независимые информационные системы.

6. Рекомендации по использованию результатов и выводов диссертации

Результаты, полученные в ходе диссертационного исследования, можно использовать в тех случаях, когда нет возможности собрать большие объемы

обучающих примеров, достаточные для применения традиционных методов математической статистики или методов обучения нейронных сетей, что встречается в промышленности, в особенности для мелкосерийных и опытных производств.

Предлагаемые методы и алгоритмы предобработки ограниченных данных рекомендуется использовать при создании систем поддержки принятия врачебных решений, поскольку зависимости состояния здоровья от медицинских показателей в виде корня принятия решений могут быть основой для построения наглядных диагностических карт, доступных для восприятия.

7. Соответствие требованиям по выполнению и оформлению результатов диссертации

Диссертация и автореферат изложены технически грамотным языком, текст хорошо структурирован и иллюстрирован, все положения аргументированы. Содержание автореферата диссертации соответствует содержанию диссертации.

Результаты исследования достаточно полно отражены в 12 публикациях (статьи и тезисы докладов), из которых три статьи – в рецензируемых научных изданиях, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, и одна статья – в издании, индексируемом в международных реферативных базах и системах цитирования Web of Science и Scopus, а также 1 статья в сборнике конференции, индексируемом в Scopus. Получен 1 патент на изобретение и 3 свидетельства о государственной регистрации программы для ЭВМ.

Основные положения и результаты диссертационной работы докладывались и обсуждались международных и всероссийских конференциях: XVIII Всероссийской школе-конференции молодых ученых «Управление большими системами» (г. Челябинск, 2022); 4th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA) (г. Липецк, 2022); XXI Всероссийской научной конференции «Нейрокомпьютеры и их применение» (г. Москва, 2023); XIX Всероссийской школе-конференции

молодых ученых «Управление большими системами» (г. Воронеж, 2023); XXXVI Международной научной конференции «Математические методы в технике и технологиях – ММТТ-36» (г. Нижний Новгород, 2023).

8 Замечания по диссертации

1) В первой главе следует пояснить более подробно почему приведен именно представляемый обзор методов анализа данных. Так, например, кластерно-регрессионный анализ также является инструментом обработки информации.

2) В третьей главе приведены интеллектуальные информационные системы. Не совсем ясно зачем это сделано с точки зрения логического перехода к разрабатываемому комплексу программного обеспечения обучения и тестирования неполносвязных нейронных сетей?

3) В четвертой главе описаны результаты кросс-валидации, при этом мало внимание уделено неполносвязным нейронным сетям. Например, интерес вызывает, как происходит зануление весов, посредством чего получаем неполносвязность в определённых узлах нейронной сети.

4) Приведено множество критериев и метрик, позволяющих оценить качество нейронных сетей. Нет пояснения, почему именно эти метрики выбраны за ключевые в оценке качества нейронных сетей.

Отмеченные недостатки не носят принципиальный характер и не влияют на общую положительную оценку диссертационной работы.

9 Заключение

Диссертационная работа Кожемякина Л.В. является законченной научно-квалификационной работой, в которой содержится научно-обоснованное решение задачи повышения качества предварительной обработки информации в условиях ограниченных наборов данных, имеющее значение для развития методов и алгоритмов системного анализа и теории управления.

Тема и содержание диссертации соответствует паспорту специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Рецензируемая диссертация отвечает требованиям Положения о порядке присуждения учёных степеней, утверждённого постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 (ред. от 25.01.2024 г.), предъявляемым к научным работам на соискание учёной степени кандидата наук, а её автор, *Кожемякин Леонид Валерьевич*, заслуживает присуждения ученой степени кандидата технических наук.

Отзыв обсужден на заседании кафедры биомедицинской инженерии от 30 августа 2024 года, протокол №1.

Отзыв составлен:

Филистом Сергеем Алексеевичем,
доктором технических наук, профессором
профессором кафедры биомедицинской инженерии
ФГБОУ ВО «Юго-Западный государственный университет»
305004, г. Курск ул. Челюскинцев, д.19, ауд. 100а,
тел.: 8 (4712) 22-26-60
e-mail: sfilist@gmail.com

Заведующий кафедрой биомедицинской инженерии
Серегин Станислав Петрович
доктор медицинских наук, профессор,
ФГБОУ ВО «Юго-Западный государственный университет»
305004, г. Курск ул. Челюскинцев, д.19, аудитория 105,
тел.: 8 (4712) 22-26-61
e-mail: kstu-bmi@yandex.ru