

ОТЗЫВ

официального оппонента

на диссертационную работу Сарина Константина Сергеевича

«Методология построения интерпретируемых нечетких

классификаторов, основанных на правилах»,

представленную на соискание ученой степени доктора технических наук по специальности 2.3.8 – Информатика и информационные процессы

1. Актуальность темы диссертационного исследования

Отсутствие понятного для человека обоснования полученного результата системами искусственного интеллекта, основанных на методах машинного обучения, вызывает опасность их использования, особенно в критически важных проблемных областях. В Национальной стратегии развития искусственного интеллекта на период до 2030 г. отмечается, что одной из причин низкого уровня доверия к современным технологиям искусственного интеллекта и препятствием для их развития является отсутствие понимания того, как искусственный интеллект достигает результатов.

Широкое применение искусственных нейронных сетей не обеспечивает высокого уровня доверия, поскольку данные модели вызывают трудности в интерпретации и объяснении полученного решения. Нечеткие системы, напротив, имеют большой потенциал для создания систем объяснимого искусственного интеллекта. База продукционных правил нечетких систем и процедура получения результата могут быть представлены на естественном языке, что позволяет доступно объяснить человеку как моделируемый процесс или явление, так и причины полученного решения.

Таким образом, диссертационная работа Сарина К.С. посвящена решению важной и актуальной проблемы – разработке и исследованию методологии построения интерпретируемых нечетких классификаторов для систем объяснимого искусственного интеллекта, предусматривающей использование предложенных в работе новых методов и алгоритмов.

2. Оценка достоверности полученных результатов и новизны диссертационного исследования

Высокая степень обоснованности научных положений, выводов и заключений в диссертационной работе подтверждается:

1) глубоким анализом современных методов построения нечетких классификаторов с определением недостатков, приводящих к низкому уровню интерпретируемости;

2) применением современных методов исследования – оптимизации, интеллектуального анализа данных, теории нечетких множеств, математической статистики;

3) корректным планированием и проведением вычислительных экспериментов;

4) практическим внедрением полученных результатов в частные и государственные организации, что подтверждено соответствующими актами.

Достоверность научных положений, выводов и заключений, содержащихся в диссертационном исследовании, оценивается как высокая. Это обусловлено следующими факторами:

1) проведением вычислительных экспериментов алгоритмов оптимизации с применением общепринятых вычислительных тестов и стандартных показателей эффективности;

2) проведением вычислительных экспериментов по построению нечетких классификаторов с применением теста перекрестной проверки и стандартных показателей точности и интерпретируемости;

3) использованием в вычислительных экспериментах достаточного количества наборов данных из различных проблемных областей и сравнение полученных результатов с известными решениями;

4) применением статистических тестов при сравнении эффективности предложенных решений с аналогами;

5) представлением результатов в публикациях рецензируемых журналов и на всероссийских и международных конференциях;

6) практическим внедрением полученных результатов.

Научную новизну диссертационной работы составляют следующие результаты.

1. Методология построения интерпретируемых нечетких классификаторов, основанных на правилах, определяющая совокупность методов анализа данных и их взаимодействие. Отличительной особенностью методологии является совместное применение методов разделения данных на обучающую и валидационную выборки, метаэвристических методов смешанной многокритериальной оптимизации для одновременной структурной и параметрической оптимизации классификаторов и стратегий

выбора классификатора из сформированного Парето-множества. Методология позволяет разрабатывать нечеткие классификаторы, сопоставимые по точности с существующими аналогами и превосходящие по показателям интерпретируемости.

2. Метаэвристический алгоритм многокритериальной оптимизации для поиска в смешанном дискретно-непрерывном пространстве. Отличительной особенностью алгоритма является совместное применение метаэвристики «кукушкин поиск» для формирования непрерывных элементов решения и генетического оператора скрещивания для формирования дискретных элементов решения. Алгоритм позволяет улучшить эффективность поиска в смешанном пространстве по сравнению с широко используемым генетическим алгоритмом с недоминируемой сортировкой NSGAII.

3. Метаэвристический алгоритм однокритериальной оптимизации на основе распределения вероятностей значений переменных для поиска в дискретном пространстве. Особенностью алгоритма является построение распределения вероятностей с помощью трансформации целевых значений решений в их весовые коэффициенты. Алгоритм позволяет улучшить эффективность поиска в бинарном пространстве по сравнению с широко используемыми генетическим алгоритмом и алгоритмом роящихся частиц.

4. Адаптированный алгоритм градиентного спуска для оптимизации параметров нечеткого классификатора. Отличительной особенностью алгоритма является представление классификатора в виде системы разделяющих функций и нормирование их значений. Алгоритм позволяет улучшить точность классификаторов, построенных метаэвристическим алгоритмом.

5. Метод разделения данных на обучающую и валидационную выборки при построении прогностических моделей машинного обучения. Метод основан на применении бинарного метаэвристического алгоритма для формирования обучающей и валидационной выборок близких по характеристикам к исходным данным. Метод позволяет улучшить точность и интерпретируемость нечетких классификаторов, построенных метаэвристическим алгоритмом смешанной многокритериальной оптимизации.

6. Методика построения нечетких классификаторов с соблюдением компромисса «точность – интерпретируемость». Отличительной особенностью методики является применение новых методов и алгоритмов, а

именно: 1) смешанной многокритериальной оптимизации с алгоритмом на основе совместного применения метаэвристики «кукушкин поиск» и генетического оператора скрещивания для одновременной структурной и параметрической оптимизации классификаторов по критериям точности и интерпретируемости, 2) метода разделения данных на обучающую и валидационную выборки дискретным алгоритмом оптимизации на основе распределения вероятностей с трансформацией целевых значений, 3) адаптированного алгоритма градиентного спуска для параметрической оптимизации классификаторов, настраивающей параметры функций принадлежности. Методика позволяет при сопоставимой точности улучшить интерпретируемость классификаторов по сравнению с известными прогностическими моделями, такими как генетические нечеткие системы FARC-HD и деревья решений CART.

Теоретическая значимость диссертационной работы состоит в решении крупной научной проблемы построения интерпретируемых прогностических моделей на основе анализа данных. Предложенную методологию построения интерпретируемых нечетких классификаторов можно квалифицировать как научное достижение в новом направлении исследований искусственного интеллекта, называемым объяснимым искусственным интеллектом. Кроме этого, метод разделения данных, разработанный Сариным К.С., может быть применен для борьбы с переобучением при построении прогностических моделей различного типа и таким образом расширяет содержание методов машинного обучения.

Практическая значимость полученных результаты диссертационного исследования подтверждается их использованием в решении конкретных практических задач. С помощью разработанной методики построения нечетких классификаторов созданы программные средства для прогнозирования эффективности реабилитации пациентов после перенесенной коронавирусной инфекции COVID-19 в Томском НИИКиФ ФФГБУ ФНКЦ МРиК ФМБА России. Прогнозирование сопровождается объяснением полученного результата, что является важным для принятия решений в области медицины. На основе алгоритма смешанной многокритериальной оптимизации «кукушкин поиск» с генетическим оператором скрещивания разработаны программные средства для синтеза цепей согласования с заданной формой амплитудно-частотной характеристики. Программные средства используются в НИИ МЭС и ООО

«Микроэлектронные системы». Алгоритм дискретной оптимизации на основе распределения вероятностей с трансформацией целевых значений внедрен в экспериментальную версию интеллектуальной САПР «50ohm Tech Circuit Studio» в компании ООО «50ом Тех.». Также результаты исследования применяются учебном процессе кафедры компьютерных систем в управлении и проектировании (КСУП) Томского государственного университета систем управления и радиоэлектроники (ТУСУР). Практическое использование подтверждается актами о внедрении. Результаты работы использованы при выполнении ряда НИР, поддержанных в рамках государственного задания Минобрнауки РФ, Российским научным фондом и Российским фондом фундаментальных исследований.

3. Оценка диссертации, степени ее завершенности, подтверждение публикаций автора

В диссертации Сарина К.С. представлено комплексное исследование, направленное на разработку методологической основы, а также новых методов и алгоритмов построения интерпретируемых моделей машинного обучения, основанных на нечетких классификаторах. Работа состоит из введения, четырех разделов, заключения, списка литературы и трех приложений. Список литературы содержит 289 наименований. Текст диссертации изложен на 260 страницах и содержит 72 рисунка и 40 таблиц. Содержание диссертации является последовательным и обоснованным, полностью отражающим проведенное исследование. Диссертация содержит достаточное количество исходных данных, имеет пояснения, рисунки, графики, подробные примеры и расчеты, написана технически квалифицированно, каждый раздел сопровождается выводами. Поставленные задачи были решены, а цель исследования достигнута. Диссертация обладает научной новизной, теоретической и практической значимостью.

Результаты диссертационного исследования прошли достаточную апробацию на конференциях всероссийского и международного уровня, опубликованы в 59 печатных работах, из которых 31 – в рекомендованных ВАК РФ периодических изданиях. Двадцать работ индексируются в международных базах научных публикаций SCOPUS и Web of Science. Получено 8 свидетельств о государственной регистрации программ для ЭВМ.

4. Соответствие автореферата основному содержанию диссертации

Автореферат имеет полную и точную информацию о содержании исследования, представлен в научном стиле с четким и логичным изложением основных положений, результатов и выводов. В нем представлены все необходимые данные для оценки научной значимости и достоверности исследования.

5. Вопросы и замечания по диссертационной работе

К диссертационной работе имеются следующие вопросы и замечания.

1. Предложенная методика построения нечетких классификаторов ограничена применением нечетких множеств с функциями принадлежности гауссова типа. Существует несколько типов стандартных функций принадлежности, например, треугольного, трапециевидного, колоколообразного типа. Автор не приводит аргументов для своего выбора.

2. В экспериментальной части исследования проводился вычислительный эксперимент по построению нечетких классификаторов с разделением пространства признаков образцов данных на три нечетких множества. Почему другие уровни детализации не использовались в эксперименте?

3. Сравнение точности классификации в экспериментальной части исследования проведено только с интерпретируемыми прогностическими моделями машинного обучения, такими как генетические нечеткие системы FARC-HD и деревья решений CART. В работе не отражены сравнения точности с результатами тестов других методов машинного обучения, например, широко используемых искусственных нейронных сетей или ансамблевых классификаторов.

4. Для оценки точности классификации автор применяет такой показатель как процент правильно классифицированных образцов. Помимо данного показателя существуют и другие, например F-мера, площадь под ROC-кривой. В диссертации не обоснован выбор показателя точности.

5. Во введении упоминается о недостатках интерпретируемости нейро-нечетких систем, в консеквентной части правил которых содержатся функции от значений признаков. Данные правила трудно интерпретировать в естественную речь. В обзорной части исследования эти системы не рассматривались. Их добавление дополнило бы представление о применении нечетких систем для классификации данных.

6. На стр. 127 диссертации в постановке задачи смешанной многокритериальной оптимизации для построения нечетких классификаторов допущена ошибка в значении верхней границы индекса i элементов вектора решений $V=(v_1, v_2, \dots, v_n)$, соответствующих непрерывным значениям. Указано $nr+1 \leq i \leq 2(nr+h_1+\dots+h_n)$, а должно быть $nr+1 \leq i \leq nr+2(h_1+\dots+h_n)$.

7. По тексту диссертации имеются опечатки, например, на стр. 97 диссертации третье предложение абзаца «Параметр нечеткого классификатора и вектор решений» начинается с «Количество параметров, настаиваемых алгоритмом...», а должно быть «Количество параметров, настраиваемых алгоритмом...».

6. Заключение

Указанные замечания не являются принципиальными, не снижают общей оценки научной и практической значимости, а также достоверности результатов диссертационного исследования.

Тема диссертации, цель и задачи работы, научная новизна, методы, исследуемые в работе, подтверждают ее соответствие паспорту научной специальности 2.3.8 – «Информатика и информационные процессы» в части следующих пунктов: п. 1 (Разработка компьютерных методов и моделей описания, оценки и оптимизации информационных процессов и ресурсов, а также средств анализа и выявления закономерностей на основе обмена информацией пользователями и возможностей и используемого программно-аппаратного обеспечения); п. 7 (Разработка методов обработки, группировки и аннотирования информации, в том числе, извлеченной из сети интернет, для систем поддержки принятия решений, интеллектуального поиска, анализа); п. 13 (Разработка и применение методов распознавания образов, кластерного анализа, нейро-сетевых и нечетких технологий, решающих правил, мягких вычислений при анализе разнородной информации в базах данных).

Диссертационная работа Сарина Константина Сергеевича на тему «Методология построения интерпретируемых нечетких классификаторов, основанных на правилах» является законченной научно-квалификационной работой, в которой разработаны теоретические положения, совокупность которых вносит существенный вклад в построение систем объяснимого искусственного интеллекта, что имеет важное научно-практическое значение для развития информатики, информационных систем и процессов.

Разработанная автором методология и в ее составе методы и алгоритмы повышают эффективность обнаружения закономерностей в данных при построении систем принятия решений, основанных на методах машинного обучения.

По своей актуальности, новизне, научно-практической значимости диссертация Сарина К.С. соответствует требованиям, предъявляемым к диссертациям на соискание ученой степени доктора наук согласно пп. 9-14 «Положения о присуждении ученых степеней», утвержденного постановлением правительства РФ от 24.09.2013 г. № 842 (в редакции от 16.10.2024 г.), а ее автор Сарин Константин Сергеевич заслуживает присуждения ему ученой степени доктора технических наук по специальности 2.3.8 – Информатика и информационные процессы.

Я, Курейчик Владимир Викторович, даю свое согласие на включение своих персональных данных в документы, связанные с работой диссертационного совета, и их дальнейшую обработку.

Официальный оппонент,
заведующий кафедрой систем
автоматизированного проектирования,
доктор технических наук (н.с. 05.13.01, 05.13.18),
профессор федерального государственного
автономного образовательного учреждения
высшего образования
«Южный федеральный университет»,
347922, г. Таганрог, пер. Некрасовский 44, кор. Г-435
тел.: (863) 438-34-51
vkur@sfedu.ru
<https://sfedu.ru/person/vkur>

«12» ноября 2025 г.



Владимир Викторович Курейчик

Подпись В.В. Курейчика заверяю
директор ИКТИБ ЮФУ, д-р техн. наук, проф.



Г. Е. Веселов