

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА
Сырямкина Владимира Ивановича на диссертацию
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

АКТУАЛЬНОСТЬ ТЕМЫ ИССЛЕДОВАНИЯ

Разработка и внедрение систем искусственного интеллекта является приоритетной задачей ближайших лет. Подавляющее большинство систем построено с помощью методов машинного обучения, т.е. без участия человека и его экспертных знаний. Построенные таким образом прогностические модели могут вызывать опасения по поводу причинения вреда, если не будет предоставлено объяснение их работы и получения результатов прогноза. Это может замедлить развитие и внедрение систем искусственного интеллекта. Широкое применение нейронных сетей не позволяет решить данную проблему, поскольку они весьма сложны для объяснения своей работы.

Нечеткие системы, являются прогностическими моделями, основанными на базе продукционных правил. Процедура получения решения в системах определяется нечетким логическим выводом. Это делает возможным извлечение знаний, описывающих закономерности в данных, а также объяснение результата полученного решения. Современные методы построения нечетких систем, осуществляющих классификацию, имеют недостатки, связанные с низким уровнем интерпретируемости получаемых моделей.

Таким образом, диссертационное исследование Сарина К.С. посвящено решению актуальной проблемы, а именно построению интерпретируемых нечетких классификаторов для систем искусственного интеллекта, позволяющих объяснить человеку результат своего решения.

АНАЛИЗ СОДЕРЖАНИЯ ДИССЕРТАЦИОННОЙ РАБОТЫ

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. Общий объем диссертации составляет 260 страниц.

Во введении обоснована актуальность темы исследования и степень ее разработанности, сформулированы цель и задачи работы, показана научная новизна, теоретическая и практическая значимость результатов. Изложены

основные положения, выносимые на защиту, и сведения об апробации работы.

В первом разделе представлен всесторонний аналитический обзор современных методов построения нечетких классификаторов на основе анализа данных. Автор выделил недостатки данных методов вследствие низкой интерпретируемости моделей. Нечеткие классификаторы типа Min-Max содержат большое количество правил, которое доходит до тысяч и приводит к невозможности интерпретировать нечеткие множества в лингвистические термины. Большое количество правил и их весовые коэффициенты усложняют интерпретируемость генетических нечетких систем. Форма функций принадлежности трехэтапного метода построения может сильно отличаться от формы равномерного разбиения. Это затрудняет интерпретацию нечетких множеств в лингвистические термины и, соответственно, представление правил на естественном языке. Кроме этого, в разделе представлена в формализованном виде проблема построения интерпретируемых нечетких классификаторов, а также используемые показатели точности и интерпретируемости.

Во втором разделе представлены новые научные результаты, направленные на решение проблемы построения интерпретируемых моделей машинного обучения, позволяющих объяснить человеку результаты своего прогноза. Предложена методология, определяющая систему методов анализа данных и их взаимодействие для построения интерпретируемых нечетких классификаторов. Методология основана на концептуальных положениях, которые учитывают выявленные недостатки методов построения в ходе аналитического обзора. Разработан алгоритм смешанной многокритериальной оптимизации, основанный на метаэвристике «кукушкин» поиск и генетическом операторе скрещивания. Разработан алгоритм дискретной оптимизации на основе распределений вероятностей с трансформацией целевых значений. Адаптирован алгоритм градиентного спуска для настройки параметров нечетких классификаторов. Разработан метод разделения данных на обучающую и валидационную выборки с применением бинарных алгоритмов оптимизации. На основе предложенной методологии, разработанных методов и алгоритмов предложена методика построения интерпретируемых нечетких классификаторов.

В третьем разделе представлена экспериментальная часть исследования. Проведены вычислительные эксперименты разработанных алгоритмов смешанной многокритериальной оптимизации и дискретного алгоритма на стандартных тестах для оценки сходимости, многообразия и стабильности их работы. Выполнено сравнение эффективности с известными аналогами. Проведены вычислительные эксперименты разработанных адаптированного алгоритма градиентного спуска, метода разделения данных и методики построения нечетких классификаторов. Эксперименты проводились на 38 общедоступных наборах данных из различных предметных областей. Результаты построения классификаторов с помощью предложенной методики сравнивались с известными интерпретируемыми

моделями деревьями решений CART и генетическими нечеткими системами FARC-HD. Сравнения проводились по критериям точности и интерпретируемости с помощью статистических тестов. Следует отметить, что построенные нечеткие классификаторы при сопоставимой точности превосходят аналоги по показателям интерпретируемости. Количество правил оказалось меньше в среднем на 39,2%, количество признаков на 8%, общее количество нечетких терминов на 49,8% по сравнению генетическими нечеткими системами FARC-HD. По сравнению с деревьями решений CART количество правил в среднем уменьшилось на 81,9%, а количество условий в правиле на 54,5%.

В четвертом разделе представлено внедрение результатов диссертационного исследования. С помощью разработанной методики построен нечеткий классификатор для прогнозирования эффективности реабилитации пациентов после перенесенной коронавирусной инфекции COVID-19. На основе классификатора созданы программные средства, осуществляющие прогнозирование эффективности реабилитации. Результаты внедрены в Томском НИИКиФ ФФГБУ ФНКЦ МРиК ФМБА России. Алгоритм смешанной многокритериальной оптимизации «кукушкин поиск» с генетическим оператором скрещивания, разработанный в диссертационной работе, применялся в программных средствах для синтеза цепей согласования в НИИ МЭС и ООО «Микроэлектронные системы». Дискретный алгоритм оптимизации на основе распределения вероятностей с трансформацией целевых значений, разработанный в диссертационной работе, внедрен в экспериментальную версию интеллектуальной САПР «50ohm Tech Circuit Studio» в ООО «50ом Тех.».

В заключении подведены итоги проделанной работы, представлены общие выводы и рекомендации. Все поставленные задачи были выполнены, а цель исследования, заключающаяся в повышении эффективности обнаружения закономерностей в данных с помощью разработки и применения методологии построения интерпретируемых нечетких классификаторов, основанных на правилах, была достигнута.

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. Тема и содержание диссертации соответствует паспорту специальности 2.3.8 – Информатика и информационные процессы.

НАУЧНАЯ НОВИЗНА

В диссертационном исследовании получены следующие новые научные результаты.

1. Предложена 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 позволила разработать нечеткие классификаторы, сопоставимые по точности с существующими аналогами и превосходящие по показателям интерпретируемости.

2. Разработан 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.

3. Разработан 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.

4. Предложен адаптированный алгоритм градиентного спуска для оптимизации параметров нечеткого классификатора, отличительной особенностью которого является представление нечеткого классификатора в виде системы разделяющих функций и нормирование их значений в целевой функции оптимизации.

5. Предложен новый метод разделения данных при построении прогностических моделей машинного обучения, основанный на применении бинарного метаэвристического алгоритма для формирования обучающей и валидационной выборок близких по характеристикам к исходным данным.

6. Разработана методика построения нечетких классификаторов с соблюдением компромисса «точность – интерпретируемость» на основе разработанных в исследовании методологии, алгоритмов оптимизации и метода разделения данных.

7. Обоснованы положения, выносимые на защиту.

ОБОСНОВАННОСТЬ И ДОСТОВЕРНОСТЬ НАУЧНЫХ ПОЛОЖЕНИЙ, ВЫВОДОВ И РЕКОМЕНДАЦИЙ, СФОРМУЛИРОВАННЫХ В ДИССЕРТАЦИИ

Высокая степень обоснованности результатов исследования не вызывает сомнения и подтверждается следующими аргументами. Проведен всесторонний аналитический обзор современных методов построения нечетких классификаторов с выделением ключевых недостатков, приводящих к низкому уровню интерпретируемости. При разработке методологии построения нечетких классификаторов, метода разделения данных и алгоритмов оптимизации автор использовал методы машинного обучения и анализа данных, теорию нечеткой логики и методы оптимизации.

В ходе проведения вычислительных экспериментов применялись стандартные тесты и показатели эффективности.

Достоверность полученных результатов исследования также не вызывает сомнений и имеет высокую степень. Это подтверждается обширными вычислительными экспериментами с применением стандартных тестов для алгоритмов оптимизации и перекрестной проверки для построения классификаторов. Оценка эффективности проводилась с помощью общепринятых показателей. Проведены сравнения с аналогами с использованием статистических тестов. Кроме этого, подтверждением достоверности результатов является их достаточное представление на конференциях и в публикациях рецензируемых изданий, а также их внедрение в практическое использование.

ТЕОРЕТИЧЕСКАЯ И ПРАКТИЧЕСКАЯ ЗНАЧИМОСТЬ

Разработанная в исследовании методология имеет высокую теоретическую значимость для развития объяснимого искусственного интеллекта. Данное направление искусственного интеллекта появилось совсем недавно и является весьма востребованным. Разработанные метаэвристические алгоритмы смешанной и дискретной оптимизации вносят существенный вклад в развитие методов оптимизации. Предложенный метод разделения данных может быть применен при построении прогностических моделей машинного обучения, как средство борьбы с переобучением.

Высокая практическая значимость работы подтверждается внедрением ее результатов в деятельность ряда организаций. Методика построения нечетких классификаторов использовалась для разработки программных средств прогнозирования эффективности реабилитации в Томском НИИКиФ ФФГБУ ФНКЦ МРиК ФМБА России. Алгоритм смешанной многокритериальной оптимизации применялся в программных средствах для синтеза цепей согласования в НИИ МЭС и ООО «Микроэлектронные системы». Дискретный алгоритм оптимизации внедрен в экспериментальную версию интеллектуальной САПР «50ohm Tech Circuit Studio» в ООО «50ом Тех.». Также результаты диссертационной работы внедрены в учебный процесс кафедры компьютерных систем в управлении и проектировании в Томском государственном университете систем управления и радиоэлектроники. Использование результатов исследования в организациях подтверждается соответствующими актами.

ПОЛНОТА ОПУБЛИКОВАНИЯ РЕЗУЛЬТАТОВ РАБОТЫ

Результаты работы достаточно апробированы на конференциях и рядом НИР, поддержанных Минобрнауки, Российским научным фондом и Российским фондом фундаментальных исследований. По теме диссертации опубликованы 59 работ, 31 публикация представлена в изданиях перечня ВАК, 20 публикаций представлены в изданиях, индексируемых в базах SCOPUS и WoS.

В автореферате изложена полная и точная информация, отражающая содержание диссертации. Представлены основные положения, результаты и выводы. Общее содержание автореферата дает полное представление о проведенном исследовании.

ЗАМЕЧАНИЯ И ВОПРОСЫ ПО ДИССЕРТАЦИОННОЙ РАБОТЕ И АВТОРЕФЕРАТУ

По диссертационной работе имеются следующие замечания.

1. В автореферате не дано пояснение, что количество условий в правиле дерева решений CART соответствует количеству терминов в правиле нечеткого классификатора при сравнении интерпретируемости по показателю среднего количества нечетких терминов в правиле. Это может быть причиной непонимания того, каким образом вычислялся данный показатель. В диссертации это пояснение есть.

2. В методе разделения данных автором были использованы при вычислении близости выборок такие характеристики как среднее значение и среднеквадратичное отклонение. Приведение примеров других характеристик расширило бы представление о применении метода.

3. В вычислительных экспериментах по построению нечетких классификаторов нечеткие множества определялись функцией принадлежности гауссова типа. Не приведены аргументы в пользу выбора данной функции.

4. В адаптированном алгоритме градиентного спуска для настройки параметров нечеткого классификатора в качестве целевой функции применялась сумма квадратов отклонений реальных выходных значений от прогнозируемых. Возможно ли применение кросс-энтропийной функции в разработанном алгоритме?

5. В описании дискретного алгоритма оптимизации на стр. 93 диссертации параметр алгоритма «вероятность мутации» обозначен p_a . В экспериментальной части исследования на стр. 171 в табл. 3.24 данный параметр обозначен как p .

Отмеченные замечания не снижают высокого уровня исследования, не влияют на полученные научные и практические результаты.

ЗАКЛЮЧЕНИЕ

Диссертация работа Сарина Константина Сергеевича на тему «Методология построения интерпретируемых нечетких классификаторов, основанных на правилах» представляет собой законченное научное исследование, содержит результаты, которые можно квалифицировать как научное достижение в области построения систем объяснимого искусственного интеллекта, и соответствует требованиям пп. 9-14 «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства Российской Федерации от 24.09.2013г. №

842 (в редакции от 16.10.2024г.), предъявляемым к диссертациям на соискание ученой степени доктора наук. Считаю, что автор диссертационного исследования Сарин Константин Сергеевич заслуживает присуждения ученой степени доктора технических наук по специальности 2.3.8 – Информатика и информационные процессы.

Даю согласие на включение своих персональных данных в документы, связанные с работой диссертационного совета, и их дальнейшую обработку.

ОФИЦИАЛЬНЫЙ ОППОНЕНТ,
профессор кафедры управления качеством ФГАОУ ВО «Национальный исследовательский Томский государственный университет»,
доктор технических наук, профессор,
заслуженный работник высшей школы РФ,
заслуженный изобретатель РФ,
лауреат премии правительства РФ в области образования

В. Сырякин

Сырямкин Владимир Иванович

24 ноября 2025 г.

Полное наименование организации: Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский Томский государственный университет»

Должность: профессор кафедры управления качеством

Ф.И.О: Сырямкин Владимир Иванович

Адрес: 634050, г. Томск, пр. Ленина, 36

Телефон: 8(3822)529-498

Эл. почта: svi_tsu@mail.ru

Подпись Сырямкина В.И. удостоверяю



Е.В. Луков