

«УТВЕРЖДАЮ»  
и.о. проректора по науке, инновациям и цифровизации  
ФГБОУ ВО «Воронежский государственный университет»,  
д.ф.м.н., доцент \_\_\_\_\_ Д.В. Костин  
\_\_\_\_\_ 2025г.



### ОТЗЫВ

ведущей организации  
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»

#### Актуальность темы исследования

В связи с интенсивным внедрением искусственного интеллекта в различные сферы человеческой деятельности возникают проблемы, связанные с возможностью информационных систем с интеллектуальной компонентой объяснять логику предлагаемых решений. Соответствующий раздел искусственного интеллекта получил название «объяснимого искусственного интеллекта», а его развитие поддерживается разработкой таких инструментов, как регрессионные модели, байесовские сети, деревья решений, а также модели, основанные на правилах. Последний тип моделей реализован в нечетких продукционных системах (или кратко – нечетких системах), которые, по сути, в общем случае являются лингвистическими аппроксиматорами. В них для входной и/или выходной переменных строятся лингвистические шкалы, а правила описывают зависимость выходной переменной от входных в форме продукционных правил (если-то правил), которые образуют основу для интерпретации результатов и, тем самым, позволяют сформировать объяснительную компоненту интеллектуальных информационных систем, основанных на знаниях. Следует заметить, что нечеткие системы реализуют принцип «серого моделирования». Модели такого типа способны генерировать полноценные естественно-языковые объяснения, а не только указывать вклад отдельных признаков и границы принятия решений. Разновидностью нечетких систем, которые успешно комбинируются с методами машинного обучения, являются нечеткие классификаторы, для них посылка продукционного правила содержит сложное лингвистическое высказывание, а заключение – метку класса. Именно лингвистическое описание входной информации и компонента знания, формализованная в виде продукционного правила, позволяют сгенерировать объяснение для предлагаемого решения. Значительный период развития нечетких систем включал исследования, посвященные обеспечению высокой точности результатов при недостаточном внимании к свойству



интерпретируемости. В связи с запросом на развитие объяснительных способностей интеллектуальных информационных систем, к которым относятся и системы на базе нечетких классификаторов, актуальной является научная задача создания методологии разработки интерпретируемых нечетких классификаторов. Таким образом, тема диссертационной работы К.С. Сарина, посвященная исследованию и построению интерпретируемых нечетких классификаторов, безусловно, является актуальной и соответствует современной проблематике искусственного интеллекта и интеллектуальных информационных систем, основанных на продукционных правилах, как форме представления знаний.

#### Структура и анализ содержания диссертационной работы

Представленная диссертация «Методология построения интерпретируемых нечетких классификаторов, основанных на правилах» изложена на 260 страницах печатного текста; состоит из введения, четырех глав, заключения, списка использованной источников, включающего 289 наименований научных трудов на русском и иностранных языках; содержит 72 рисунка и 40 таблиц.

Введение содержит стандартные для диссертации разделы.

В *первой* главе введено понятие нечеткой системы и нечеткого классификатора, как ее частного случая; представлен обзор распространенных нечетких классификаторов со свойством интерпретируемости (FMM и его модификации; генетический нечеткий классификатор с различными вариантами реализации и многоцелевые генетические нечеткие системы) и выявлены их недостатки, что позволило определить направления совершенствования нечетких классификаторов относительно свойства интерпретируемости; изложен трехэтапный метод построения нечетких классификаторов, который допускает различные варианты реализации, в том числе с привлечением процедур кластеризации для определения структуры нечеткого классификатора и генетической настройки параметров правил. Также в этой главе вводятся неформальное и формальное понятия интерпретируемости. Интерпретируемость можно ассоциировать с возможностью объяснить работу модели на естественном языке. К показателям интерпретируемости относятся: количество правил, количество признаков, среднее количество нечетких терминов в правилах, общее количество нечетких терминов в базе правил. Сформулирована задача построения нечеткого классификатора как задача многоцелевой оптимизации, где в качестве целевых функций выступают показатели точности и интерпретируемости. Материал первой главы является теоретической основой данного исследования.

Во *второй* главе представлена методология построения нечетких классификаторов в форме системы взаимосвязанных моделей, методов и



алгоритмов анализа данных, которая с учетом недостатков существующих подходов нацелена на обеспечение свойства интерпретируемости, а также обоснованного компромисса между интерпретируемостью и точностью. Основу методологии составляет оптимизационный подход к решению задач структурной и параметрической оптимизации нечеткого классификатора, при этом его особенностями являются: использование метаэвристических алгоритмов; смешанная оптимизация, осуществляемая одновременно по дискретным и непрерывным переменным; использование оптимизационных методов, основанных на производных для дообучения модели, поскольку они осуществляют поиск локальных экстремумов с более высокой точностью; специальный подход к формированию обучающего множества для настройки параметров нечеткого классификатора на основе метода разделения данных.

Третья глава посвящена описанию проведенных вычислительных экспериментов и сравнению полученных результатов с аналогами. Оценка эффективности разработанных алгоритмов оптимизации, метода разделения данных и методики построения нечетких классификаторов проведена с помощью общепринятых тестов и на общедоступных наборах данных. Сравнительный анализ проведен с помощью статистических тестов, результаты анализа подтвердили статистически значимое преимущество разработанных методов и алгоритмов.

В четвертой главе дано описание практического применения результатов диссертационной работы при решении задач прогнозирования эффективности реабилитации пациентов после перенесенной коронавирусной инфекции COVID-19, синтеза согласующих цепей на основе смешанной многокритериальной оптимизации и синтеза секции ступенчатого аттенюатора с помощью дискретного алгоритма оптимизации.

Заключение содержит основные результаты диссертационного исследования и выводы.

#### Научная новизна полученных результатов, выводов и рекомендаций

В диссертации получены следующие результаты, характеризующиеся научной новизной:

– основы методологии построения интерпретируемых нечетких классификаторов, основанных на если-то правилах, в форме системы взаимосвязанных методов анализа данных и алгоритмов параметрической и структурной оптимизации базы знаний нечеткого классификатора с целью повышения уровня интерпретируемости и обеспечения точности, при этом отличительными особенностями методологии являются: необходимость тщательного подхода для формирования обучающего множества; использование методов смешанной многокритериальной оптимизации с переменными



дискретного и непрерывного типов; возможность дообучения моделей с применением методов оптимизации, основанных на производных; использование различных стратегий выбора на фронте Парето для выбора модели нечеткого классификатора;

- *метаэвристический алгоритм многоцелевой оптимизации* для случая, когда целевые функции заданы в дискретно-непрерывном пространстве, отличающийся совместным применением метаэвристического алгоритма «кукушкин поиск» и генетического оператора скрещивания, что позволило существенно улучшить эффективность поиска по сравнению с широко известным алгоритмом многоцелевой оптимизации NSGAII;

- *дискретный метаэвристический алгоритм оптимизации*, отличительной особенностью которого является определение распределения вероятностей с помощью трансформации целевых значений элементов решений в их весовые коэффициенты, что позволило повысить эффективность поиска в бинарном пространстве по сравнению с генетическим алгоритмом и алгоритмом роящихся частиц (PSO);

- *адаптированный алгоритм градиентного спуска* для решения задачи параметрической оптимизации функций принадлежности лингвистических шкал входных переменных, отличающийся применением целевой функции с нормированными значениями разделяющих функций и позволяющий улучшить точность нечеткого классификатора;

- *метод формирования обучающих и валидационных данных* с сохранением близости полученных выборок к исходной без потери важной для обучения и валидации информации с учетом статистических характеристик, позволяющий улучшить точность и интерпретируемость нечетких классификаторов, построенных метаэвристическим алгоритмом смешанной многокритериальной оптимизации.

- *методика построения нечетких классификаторов*, включающая комплекс разработанных алгоритмов, базирующаяся на предложенной методологии и обеспечивающая, тем самым, компромисс между критериями точности и интерпретируемости, что позволило при сопоставимой точности улучшить интерпретируемость нечетких классификаторов по сравнению с известными моделями, такими как генетический нечеткий классификатор FARC-HD и алгоритм CART. Методика построения нечетких классификаторов является инструментом извлечения знаний из данных и интерпретации результатов, облегчающих понимание и объяснение процессов принятия решений, связанных с классификацией объектов.

Полученные в диссертационной работе результаты являются новыми, теоретически обоснованными, а их преимущества перед существующими



решениями в этой области – экспериментально подтвержденными. Совокупность научных результатов можно квалифицировать как научное достижение в области построения систем объяснимого искусственного интеллекта

Все результаты работы, составляющие научную новизну и сформулированные в положениях, выносимых на защиту, получены автором лично.

Степень обоснованности научных положений, выводов и рекомендаций,  
сформулированных в диссертации

Представленные в работе научные положения, выводы и рекомендации следует признать обоснованными. Автор проявляет широкий кругозор в предметной области, корректно использует современные научные методы исследования и информационные технологии. При получении практически каждого теоретического или экспериментального вывода проводился сравнительный анализ с существующими решениями, подходами и алгоритмами. Каждый вывод сопровождается четкой качественной и количественной интерпретацией. Обоснованность подтверждается результатами вычислительных экспериментов и практическим внедрением разработанных решений.

Достоверность результатов диссертации подтверждается корректным применением методов оптимизации, нечеткого моделирования и кластерного анализа; результатами применения взаимно дополняющих теоретических и практических методов исследования, их согласованностью с известными результатами других исследователей; использованием общепринятых тестов и общедоступных наборов данных.

Экспериментальная часть исследования имеет существенный объем. В каждом случае определена методика проведения эксперимента. Оценка алгоритмов оптимизации проводилась с помощью общепринятых тестов, разработанных для нахождения оптимума унимодальных и мультимодальных тестовых функций. Для исследования предложенных методов и оценки значимости различий применялся статистический анализ. Построение нечетких классификаторов для оценки эффективности предлагаемой методики осуществлялась на общедоступных наборах данных из репозитория KEEL (<http://keel.es>). Сравнение с известными интерпретируемыми моделями машинного обучения, такими как нечеткие классификаторы FARC-HD и деревья решений CART, показало преимущества в точности и интерпретируемости нечетких классификаторов, построенных по предложенной автором методике.

Для моделирования и постановки экспериментов использовалось лицензионное программное обеспечение. На модули и отдельные программы получены Свидетельства о государственной регистрации программ.



### Значимость результатов исследования для науки и практики

Теоретическая значимость полученных автором результатов заключается в развитии технологии построения интерпретируемых систем интеллектуального анализа данных. Методика построения нечетких классификаторов является инструментом обнаружения закономерностей в данных с возможностью объяснения полученных результатов на языке, близком к естественному. Алгоритмы смешанной многокритериальной оптимизации и дискретной оптимизации могут применяться для решения задач глобальной параметрической оптимизации и комбинаторной оптимизации. Метод формирования обучающих и валидационных данных может быть использован как инструмент для устранения переобучения и выбора подходящей по сложности структуры при построении моделей машинного обучения различных типов.

Практическая ценность работы заключается в создании и внедрении инструментария, позволяющего решить одну из ключевых проблем разработки информационных систем, основанных на знаниях – проблему оптимизации базы знаний, в большой степени влияющую на качество решения прикладных задач. Алгоритмы и программные средства, разработанные в диссертации, использовались для синтеза согласующих цепей с заданной формой амплитудно-частотной характеристики; в экспериментальной версии интеллектуальной САПР; для построения интерпретируемых моделей прогнозирования эффективности реабилитации пациентов после перенесенной коронавирусной инфекции COVID-19. Все внедрения подтверждаются соответствующими актами, копии которых приводятся в Приложении.

Диссертационная работа поддержана грантами 5 научно-исследовательских проектов (НИР, РНФ, РФФИ).

Результаты диссертационной работы рекомендуются для использования в научно-исследовательских, проектных и промышленных организациях, занимающихся исследованиями в области разработки интеллектуальных информационных систем и экспертных систем различного назначения, ядром которых являются нечеткий классификатор и метаэвристические методы оптимизации (например, АО «НПП «Исток» им. Шокина» (г. Фрязино), АО «ОКБ-Планета» (г. Великий Новгород), АО «Светлана-Рост» (г. Санкт-Петербург), УПКБ «Деталь» (г. Каменск-Уральский), АО «Микрон» (г. Зеленоград) и др.).

Отдельные разделы работы целесообразно использовать в учебном процессе вузов, в которых осуществляется подготовка по направлениям, связанным с искусственным интеллектом и машинным обучением: Воронежский государственный университет, Южный федеральный университет, Санкт-Петербургский государственный университет, Московский физико-технический институт, университет «Иннополис», Национальный исследовательский университет ИТМО и др.



Соответствие критериям, которым должны отвечать диссертации  
на соискание ученых степеней

Диссертация Сарина К.С. является завершенной научно-квалификационной работой, в которой содержится решение важной для инженерии знаний и разработки систем объяснимого искусственного интеллекта научной задачи совершенствования подходов к генерации баз знаний нечетких классификаторов с улучшенными свойствами по точности и интерпретируемости (п.9). Цель диссертационной работы, заключающаяся в повышении эффективности обнаружения закономерностей в данных с помощью разработки и применения методологии построения интерпретируемых нечетких классификаторов, основанных на правилах, была достигнута, а все поставленные задачи были выполнены.

Диссертация обладает внутренним единством, логически правильно выстроена, содержит новые научные результаты, полученные лично автором. Предложенные соискателем решения аргументированы и оценены по сравнению с известными решениями (п. 10).

Все научные результаты диссертации опубликованы в рецензируемых научных изданиях. Соискатель имеет 59 научных работ, в том числе 31 – в рекомендованных ВАК РФ периодических изданиях; 20 – в международных базах научных публикаций Scopus и WoS. Получены восемь свидетельств о государственной регистрации программ для ЭВМ. Качество и количество публикаций соискателя по теме диссертации удовлетворяют требованиям ВАК (пп. 11, 13).

В диссертации имеются обязательные ссылки на источники заимствования материалов. В диссертации и автореферате указан личный вклад соискателя в работы, выполненные в соавторстве. (п. 14).

Автореферат в полной мере отражает содержание и основные положения диссертации.

Замечания по диссертационной работе

1) Кажется, что введение в нечеткие системы очень краткое. Это инструмент, который работает со знаниями в форме продукционных правил. Можно было бы побольше написать про проблему формирования базы знаний (базы правил) для нечетких систем. Для базы знаний актуальной является проблема структурной и параметрической оптимизации (оптимизируется количество правил, количество термов в лингвистических шкалах, параметры функций принадлежности). Нужно было бы четко сформулировать отношение к данной проблеме, упомянуть ее.



2) Формальное понятие интерпретируемости автор характеризует несколькими количественными показателями (количество правил, количество терминов в базе правил и др.), но кажется, что они являются зависимыми.

3) Признаки, используемые в классификаторе, могут быть устроены по-разному: 1) предпочтительно иметь большее значение; 2) предпочтительно иметь меньшее значение; 3) предпочтительно иметь некоторое заданное значение, поэтому функции нормирования должны эти принципы. Однако этот факт не учитывается при нормировании данных (стр. 119).

4) При разбиении пространства входных признаков автор использует гауссовы функции принадлежности и вводит дополнительные параметры, чтобы определить боковое смещение и смещение охвата функции принадлежности. Однако можно было бы использовать более обобщенное представление гауссовой функции принадлежности и настраивать алгоритмом ее параметры, которые «отвечают» и за смещение, и за деформацию функции принадлежности.

5) Целесообразно указать, в каком диапазоне изменяется классовый вес правила, и равна ли сумма весов 1?

6) В таблицах, содержащих результатов экспериментов, следовало бы выделить лучшие результаты.

7) Не совсем понятно, почему из множества эвристических алгоритмов выбран алгоритм «кукушкин поиск»?

8) Как следует из текста диссертации существуют определенные ограничения на предложенную методику (например, лингвистические шкалы фиксированные, содержат одинаковое число термов для всех входных переменных; используется только смещение и равномерная деформация функций принадлежности и др.). Можно ли предположить, что в рамках методологии каждая входная переменная имеет индивидуальную лингвистическую шкалу с произвольным числом термов, а в результате оптимизации (настройки параметров) может получиться неравномерная лингвистическая шкала. Тогда методология (методика) будет носить более универсальный характер. Следовало бы привести рассуждения по поводу обобщения предложенной методики.

Перечисленные замечания не снижают положительного впечатления от представленной работы, не снижают ее теоретическую и практическую значимость.

### Заключение

Содержание диссертации и автореферата свидетельствуют о высоком уровне проведенных исследований, отличающихся глубиной проработки, оригинальностью подходов и убедительностью результатов. Заявленные научные результаты обладают научной новизной и вносят существенный вклад в развитие методологии построения интеллектуальных систем. В исследовании



представлены новые решения, совокупность которых можно квалифицировать как научное достижение в области построения систем объяснимого искусственного интеллекта. Диссертация Сарина Константина Сергеевича по содержанию и представленным результатам соответствует паспорту специальности 2.3.8 – Информатика и информационные процессы. Научные и практические результаты, полученные при проведении диссертационного исследования, и полнота их освещения в публикациях соискателя позволяют считать, что диссертация является завершенной научно-квалификационной работой, удовлетворяет требованиям пп. 9-14 «Положения о присуждения ученых степеней», предъявляемым к докторским диссертациям и утвержденным постановлением правительства РФ от 24.09.2013г. № 842 (в редакции от 16.10.2024г.), а ее автор Сарин Константин Сергеевич заслуживает присуждения ученой степени доктора технических наук по специальности 2.3.8 – Информатика и информационные процессы.

Отзыв составлен:

Доцент кафедры вычислительной математики  
и прикладных информационных технологий,  
к.ф.-м.н., доцент



Медведев Сергей Николаевич

Утвержден:

на заседании вычислительной математики  
и прикладных информационных технологий,  
протокол № 4 от 21 ноября 2025 г.

Зав. кафедрой вычислительной математики  
и прикладных информационных технологий  
ФГБОУ ВО «Воронежский государственный  
университет», д.т.н., профессор



Леденева Татьяна Михайловна

Федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет»

Почтовый адрес: 394018, г. Воронеж, Университетская пл., д. 1

Телефон: +7 (473) 220-75-21

