

Адаптированный алгоритм градиентного спуска для настройки параметров нечеткого классификатора

Сарин Константин Сергеевич

Томский государственный университет систем управления и радиоэлектроники,
Россия, Томск, sarin.konstantin@mail.ru

Аннотация. Использование систем искусственного интеллекта для критически важных сфер жизнедеятельности человека требует доверия к полученному системой результату. Важным для обеспечения доверия является объяснение полученного решения. Нечеткие системы обладают свойством объяснимости благодаря наличию базы продукционных правил на естественном языке. Настоящая работа посвящена модификации алгоритма градиентного спуска для настройки параметров нечетких классификаторов. Эксперименты на 38 наборах данных показали, что применение алгоритма для классификаторов, построенных метаэвристическим алгоритмом, статистически значимо увеличивает точность классификации.

Ключевые слова: нечеткие системы, классификация, анализ данных, алгоритмы оптимизации

Цитирование: Сарин К.С. Адаптированный алгоритм градиентного спуска для настройки параметров нечеткого классификатора / К.С. Сарин // Информационные и математические технологии в науке и управлении, 2025. – № 2 (38). – С. 18-31. – DOI:10.25729/ESI.2025.38.2.002.

Введение. Системы искусственного интеллекта (СИИ) нашли широкое применение во многих сферах жизнедеятельности человека, в том числе и таких критически важных, как здравоохранение, энергетика и оборона [1]. Поскольку решения, принимаемые такими системами, влияют на жизни людей, появляется необходимость понимания того, как были получены такие решения [2, 3]. Отсутствие такой возможности может повлиять на уровень доверия человека к результатам решений этой системы и может вызвать опасения по поводу причинения вреда для себя и общества в целом [4, 5]. Это может замедлить общественное принятие и внедрение СИИ. В Национальной стратегии развития искусственного интеллекта на период до 2030 года, утвержденной указом Президента Российской Федерации от 10.10.2019 № 490 (ред. от 15.02.2024), отмечается, что отсутствие понимания того, как искусственный интеллект достигает результатов, является одной из причин низкого уровня доверия к современным технологиям искусственного интеллекта и может стать препятствием для их развития.

Решение указанной выше проблемы предлагается осуществлять с помощью объяснимого искусственного интеллекта (ОИИ). СИИ обладает свойством объяснимости, если способна предоставлять доказательства или обоснование человеку в поддержку своего решения [6, 7]. Взаимодействие человека и СИИ играет ключевую роль в ОИИ. Полученные системой знания должны использоваться для предоставления объяснений; кроме того, пользователи должны иметь возможность оспаривать и задавать вопросы системам, чтобы понять обоснование решения в соответствии со своими конкретными потребностями [8].

В последние годы произошел рост систем поддержки принятия решений, основанных на нейронных сетях глубокого обучения. Эмпирический успех моделей глубокого обучения обусловлен сочетанием эффективных алгоритмов обучения и их огромного параметрического пространства, которое может включать сотни слоев и миллионы параметров [9]. Это является причиной того, что такие СИИ рассматриваются, как сложные модели «черного ящика» [10], исключающие прозрачность или поиск понимания механизма, с помощью которого работает модель [11].

Нечеткие системы имеют большой потенциал в разработке решений ОИИ. Благодаря наличию базы нечетких правил они способны выражать знания в ориентированной на

человека форме, используя термины естественного языка [12]. Такая возможность позволяет предоставить понятные объяснения используемых системой знаний, а также детали, касающиеся процесса вывода при принятии решений [13]. С помощью теории нечеткой логики, лежащей в основе таких систем, была установлена связь между числовыми решениями и правилами на естественном языке. Таким образом, численные операции, осуществляющие процедуру принятия решения в нечетких системах, интерпретируются логическими рассуждениями во вполне понятной человеку лингвистической форме. Это побудило использовать нечеткие методы для разработки ОИИ [14].

Нечеткие классификаторы являются нечеткими системами, осуществляющими процесс классификации или нахождения класса объекта по значениям его признаков [15]. Продукционные правила таких систем связывают значения признаков объекта, представленных в лингвистической форме, с их классом. Построение нечетких классификаторов на основе анализа данных методами машинного обучения рассматривается, как проблема оптимизации или поиска [16], поэтому для их построения применяются алгоритмы оптимизации. Детерминированные алгоритмы оптимизации применялись для настройки параметров нечетких аппроксиматоров. Алгоритм градиентного спуска применялся для настройки параметров адаптивных нейро-нечетких систем вывода (ANFIS), основанных на моделях Такаги-Сугено и используемых для решения задач аппроксимации [17]. Данные системы также применялись и для классификации, но их объяснимость для человека весьма затруднительна, поскольку консеквентная часть нечетких правил таких систем содержит не метку класса объекта, а линейную функцию от значений входных признаков [18-20]. Для построения нечетких классификаторов широко применялись эвристические методы оптимизации, основанные на эволюционных и роевых метаэвристических алгоритмах [16, 21-23].

Предпочтение в выборе метаэвристических алгоритмов для построения нечетких классификаторов заключается в их способности осуществлять поиск глобального оптимума, в отличие от детерминированных, которые сосредоточены на локальном поиске. Также данные алгоритмы не требуют анализа целевой функции и ее дифференцируемости, что является затруднительным для нечетких классификаторов. Фактически целевая функция представляется «черным ящиком», для которого определяются только входные и выходные значения [24], но в детерминированных алгоритмах, основанных на дифференцировании, точность поиска локального оптимума в окрестности заданной точки значительно выше. Поэтому комбинированный поиск на основе метаэвристического и детерминированного алгоритмов потенциально может повысить эффективность поиска глобального оптимума.

В настоящей работе предлагается адаптировать алгоритм градиентного спуска для настройки параметров нечетких классификаторов и применить его после построения классификаторов метаэвристическим алгоритмом. Исходные значения параметров, найденные с помощью метаэвристики, могут быть улучшены более точным локальным поиском градиентного спуска. Такое дообучение параметров классификаторов может повысить точность их прогнозов.

Цель настоящей работы заключается в повышении точности нечетких классификаторов, построенных метаэвристическими алгоритмами, с помощью применения адаптированного алгоритма градиентного спуска для настройки их параметров.

Основной научный вклад представлен следующими результатами.

1. Адаптирован алгоритм градиентного спуска для настройки параметров функций принадлежности нечетких классификаторов.

2. Эмпирически доказана эффективность предложенного алгоритма при построении классификаторов на 38 наборах данных из различных предметных областей. Статистический тест Уилкоксона показал значимое повышение точности классификации после применения алгоритма к классификаторам, построенным метаэвристическим алгоритмом смешанной многокритериальной оптимизации.

Остальная часть статьи организована следующим образом. В п.2 рассмотрены архитектура нечеткого классификатора и его построение на основе анализа данных с помощью алгоритма смешанной многокритериальной оптимизации. Предложен адаптированный алгоритм градиентного спуска для настройки параметров. В п.3 описана экспериментальная часть исследования для оценки эффективности предлагаемого алгоритма. В заключении сделаны выводы о проделанной работе.

2. Построение нечеткого классификатора. Согласно каноническому описанию, представленному Дудой и Хартом [15], классификатор – это устройство, определяющее класс объекта $\mathbf{x} = (x_1, \dots, x_n)$ среди множества классов $\{\omega_1, \omega_2, \dots, \omega_M\}$ посредством системы разделяющих функций $g_k(\mathbf{x})$, $k = 1, \dots, M$. Каждая функция связана с конкретным классом и возвращает значение, которое характеризует степень уверенности в том, что объект принадлежит данному классу. Для того, чтобы определить класс объекта, вычисляются значения разделяющих функций. Классом объекта является класс функции с максимальным значением.

Предположим, что имеется m обучающих (т.е. с известным классом) образцов данных $\mathbf{x}_p = (x_{p1}, \dots, x_{pn})$, $p = 1, 2, \dots, m$.

Каждый образец принадлежит одному из M классов $\{\omega_1, \omega_2, \dots, \omega_M\}$ в n -мерном пространстве признаков, где x_{pi} – значение i -й переменной p -го образца ($i = 1, \dots, n$). В настоящем разделе представлены алгоритмы, которые на основе обучающих образцов данных осуществляют построение нечеткого классификатора. По своей сути, нечеткий классификатор – это математическая модель на основе теории нечеткой логики, определяющая зависимость между значениями признаков объекта и его классом.

2.1. Архитектура нечеткого классификатора. Архитектура нечеткого классификатора, используемого в работе, впервые предложена в [25] и впоследствии нашла широкое применение во многих исследованиях. Нечеткий классификатор состоит из продукционных правил «ЕСЛИ-ТО» с нечеткими терминами в антецедентной части, оценивающими значения признаков образца, и метками классов в консеквентной части, определяющими его класс:

Правило R_j : ЕСЛИ $x_1 = A_{1j}$ И ... И $x_n = A_{nj}$ ТО class = c_j , $j = 1, \dots, R$,

где R – количество правил; $\mathbf{x} = (x_1, \dots, x_n)$ – вектор входных переменных, соответствующий признакам классифицируемого образца $\mathbf{x}$; A_{ij} – нечёткий термин, характеризующий i -ю переменную в j -ом правиле ($i = 1, \dots, n$); c_j – метка класса в правиле.

Формализация нечетких терминов осуществляется с помощью функций принадлежности, т.е. с каждым термином A_{ij} ассоциируется функция принадлежности $\mu_{ij}(x_i)$, определенная на множестве значений x_i . В работе использовались функции принадлежности гауссова типа, содержащие два параметра: среднее значение c и стандартное отклонение σ :

$$\mu(x) = e^{-\frac{(x-c)^2}{2\sigma^2}}.$$

Если заданы база правил и функции принадлежности, соответствующие нечетким терминам в правилах, то определение класса нового образца $\mathbf{x}$ среди классов $\{\omega_1, \omega_2, \dots, \omega_M\}$ осуществляется с помощью процедуры нечеткого вывода. В выводе для логической операции И используется t -норма в виде произведения, а агрегация правил проводится суммированием.

На основе этих операций вычисляются значения разделяющих функций образца для каждого класса:

$$g_k(\mathbf{x}) = \sum_{\substack{j=1,\dots,R, \\ c_j=\omega_k}} \prod_{i=1}^n \mu_{ij}(x_i), \quad (1)$$

где $k = 1, \dots, M$. Затем определяется класс образца, который соответствует классу разделяющей функции с максимальным значением:

$$\mathbf{x} \in \omega_k, k = \arg \max_{k=1,\dots,M} (g_k(\mathbf{x})).$$

2.2. Построение классификатора метаэвристическим алгоритмом. В исследовании нечеткий классификатор создается с помощью алгоритма смешанной многокритериальной оптимизации «кукушкин поиск» с генетическим оператором скрещивания [26]. Этот алгоритм показал высокую эффективность на комплексных оценках сходимости и многообразия и позволяет проводить поиск в смешанном дискретно-непрерывном пространстве, в котором присутствуют как дискретные, так и вещественные переменные.

Перед построением классификатора необходимо задать количество лингвистических терминов, связанных с каждой переменной, а также функции принадлежности, соответствующие этим терминам. Для повышения уровня интерпретируемости, предлагается использовать равномерное разбиение каждой переменной на нечеткие множества и при построении классификатора осуществлять малые изменения параметров функций принадлежности. В [27] для функций принадлежности равномерного разбиения вводится параметр α для бокового смещения функции. Значение «1» данного параметра соответствует увеличению среднего значения на величину, соответствующую расстоянию между центрами соседних функций принадлежности равномерного разбиения. В настоящей работе помимо бокового смещения предлагается использовать смещение охвата β для изменения среднеквадратичного отклонения. Допустим, что пространство признака равномерно разбито на H нечетких множеств. Значение «1» данного параметра соответствует увеличению среднеквадратичного отклонения на величину, равную разности между значениями среднеквадратичных отклонений функций принадлежности при разбиении на $H-1$ и H нечетких множеств. Чтобы получить значения параметров функций принадлежности с учетом введенных смещений, необходимо провести следующие преобразования:

$$s_{ij} = \frac{x_i^{\max} - x_i^{\min}}{H-1} (j-1 + \alpha_{ij}), \sigma_{ij} = \frac{(x_i^{\max} - x_i^{\min})}{3(H-1)} \left(1 + \frac{\beta_{ij}}{H} \right), \quad (2)$$

где s_{ij} , b_{ij} – среднее значение и среднеквадратичное отклонение функции принадлежности i -й переменной j -го термина; α_{ij} , β_{ij} – соответствующие боковое смещение и смещение охвата.

Параметры α и β при малых значениях инициируют малые сдвиги, что не существенно отражается на потере интерпретируемости. В настоящей работе оба смещения изменяются в диапазоне $[-0,3, 0,3]$. С помощью выражений (2) они преобразуются в параметры функции принадлежности. На рисунке 1 показано равномерное разбиение переменной на три нечетких термина «Малое», «Большое» и «Среднее», а также боковые смещения функции принадлежности нечеткого множества «Среднее» для $\alpha = -0,3$ и $\alpha = 0,3$ и смещения охвата для $\beta = -0,3$ и $\beta = 0,3$. Для того, чтобы повысить интерпретируемость, крайние слева и справа функции равномерного разбиения не имеют бокового смещения, а только смещение охвата. Таким образом, при разбиении на три нечетких множества боковое смещение при построении классификатора будет иметь только функция принадлежности термина «Среднее».

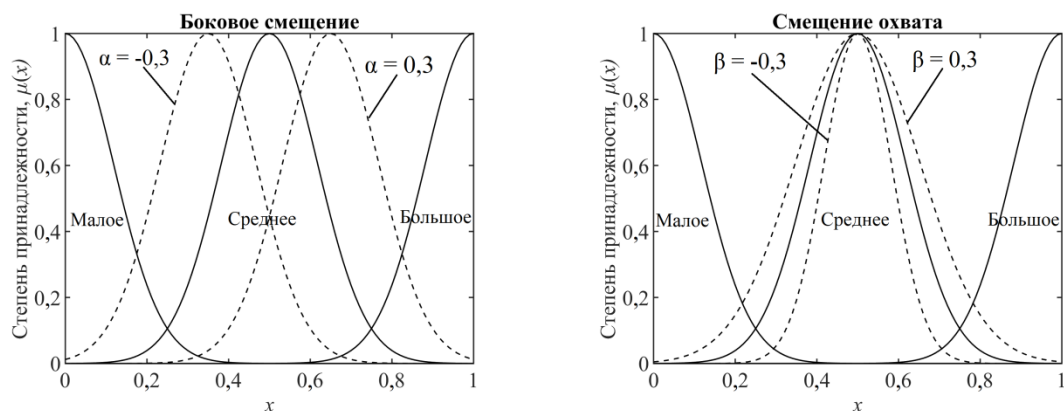


Рис. 1. Функции принадлежности с боковым смещением и смещением охвата

С помощью метаэвристического алгоритма смешанной многокритериальной оптимизации осуществляется построение классификатора. Целочисленная часть вектора решений алгоритма кодирует antecedentную часть базы правил классификатора путем сопоставления переменной номера нечеткого термина, образованного равномерным разбиением. Вещественная часть содержит боковое смещение и смещение охвата для функций принадлежности каждого термина разбиения. Консеквентная часть правила, определяющая класс, не присутствует в векторе поиска, поскольку может быть вычислена на основе функций принадлежности antecedentной части и обучающих образцов данных с помощью эвристики, предложенной в [28, 29]. В качестве целевых значений оптимизации используются показатель точности – процент правильно классифицированных образцов – и показатели интерпретируемости – количество правил и общее количество терминов в базе правил. Поскольку алгоритм возвращает множество классификаторов, образующих недоминируемые по Парето решения, из этого множества выбирается классификатор с лучшей точностью на валидационной выборке.

2.3. Адаптированный алгоритм градиентного спуска. Задача алгоритма градиентного спуска заключается в нахождении параметров функций принадлежности, увеличивающих точность классификации. Обычно точность классификации оценивается с помощью показателей, не пригодных для дифференцирования, например, таких, как процент правильно классифицированных образцов или f-мера. Применение градиентных методов оптимизации, в том числе и алгоритма градиентного спуска, требует дифференцируемость целевой функции. Для нахождения параметров нейронных сетей с помощью варианта алгоритма градиентного спуска, называемого обратным распространением ошибки, в качестве целевой функции обычно используют сумму квадратов отклонений значений разделяющих функций и их ожидаемых значений. Эта функция должна быть минимизирована:

$$E = \frac{1}{2} \sum_{p=1}^m \sum_{k=1}^M (g_k(\mathbf{x}_p) - y_{pk})^2, \quad (3)$$

где $\mathbf{x}_p = (x_{p1}, \dots, x_{pn})$ – p -й образец обучающих данных, y_{pk} – ожидаемое значение разделяющей функции k -го класса для p -го образца. Если образец принадлежит классу ω_h , т.е. $\mathbf{x}_p \in \omega_h$, то ожидаемые значения разделяющих функций данного образца определяются следующим образом:

$$y_{pk} = \begin{cases} 1, & \text{если } k = h \\ 0, & \text{иначе} \end{cases}, k = 1, \dots, M.$$

Множитель $1/2$ в (3) используется для того, чтобы избавиться от множителя 2 после дифференцирования функции, на значение оптимального решения он не влияет.

Использование целевой функции (3) для нахождения параметров нечеткого классификатора сталкивается со следующей проблемой. В результате агрегации правил значение g_k может быть больше 1, хотя ожидаемые значения ограничены диапазоном $[0, 1]$. Чтобы избавиться от вышеприведенной проблемы при нахождении параметров нечеткого классификатора, предлагается изменить целевую функцию оптимизации с помощью нормирования значений разделяющих функций путем деления на их сумму:

$$E = \frac{1}{2} \sum_{p=1}^m \sum_{k=1}^M \left(\frac{g_k(\mathbf{x}_p)}{\sum_{k=1}^M g_k(\mathbf{x}_p)} - y_{pk} \right)^2. \quad (4)$$

В таком случае они будут ограничены диапазоном $[0, 1]$.

Основой метода градиентного спуска является понятие градиента целевой функции. Градиент представляет собой вектор частных производных функции по ее переменным и показывает направление в сторону ее наискорейшего возрастания. Далее будут найдены формулы для вычисления частных производных по параметрам s и σ классификатора. Применяя правила дифференцирования сложной, степенной и дробной функции по параметру b к (3), получается следующее значение частной производной:

$$E'_b = \sum_{p=1}^m \sum_{k=1}^M \left(\frac{g_k(\mathbf{x}_p)}{\sum_{k=1}^M g_k(\mathbf{x}_p)} - y_{pk} \right) \frac{g'_{kb}(\mathbf{x}_p) \left(\sum_{k=1}^M g_k(\mathbf{x}_p) \right) - \left(\sum_{k=1}^M g'_{kb}(\mathbf{x}_p) \right) g_k(\mathbf{x}_p)}{\left(\sum_{k=1}^M g_k(\mathbf{x}_p) \right)^2}. \quad (5)$$

Чтобы рассчитать данное значение, необходимо вычислить значения производных разделяющих функций g'_{kb} , где $k = 1, \dots, M$. Для этого обозначим множество номеров правил k -го класса, содержащих в antecedентной части функцию принадлежности с параметром b , как R^b_k , а номер соответствующей переменной в правиле как d . Значение производной разделяющей функции по параметру среднего s функции принадлежности согласно (1) определяется следующим образом:

$$g'_{ks}(\mathbf{x}_p) = \left(\sum_{\substack{j=1, \dots, R, \\ c_j = \omega_k}}^n \prod_{i=1}^n \mu_{ij}(x_{pi}) \right)'_s = \frac{x_{pd} - s}{\sigma^2} \sum_{j \in R^s_k} \prod_{i=1}^n e^{-\frac{(x_{pi} - s_{ij})^2}{2\sigma_{ij}^2}} \quad (6)$$

Аналогично находится производная по параметру среднеквадратичного отклонения σ :

$$g'_{k\sigma}(\mathbf{x}_p) = \frac{(x_{pd} - s)^2}{\sigma^3} \sum_{j \in R^s_k} \prod_{i=1}^n e^{-\frac{(x_{pi} - s_{ij})^2}{2\sigma_{ij}^2}} = \frac{x_{pd} - s}{\sigma} g'_{ks}(\mathbf{x}_p) \quad (7)$$

Блок-схема алгоритма градиентного спуска для оптимизации параметров нечеткого классификатора представлена на рисунке 2. Обозначим параметры вектором $\mathbf{P} = (p_1, p_2, \dots, p_n)$, каждый элемент которого представляет параметры среднего значения или среднеквадратичного отклонения функций принадлежности. Для поиска минимума целевой функции E (5) необходимо изменять параметры нечеткого классификатора в направлении, противоположном направлению градиента, причем при изменении параметры не должны выйти за границы их значений. Границы являются параметрами алгоритма и задаются до выполнения его работы. На блок-схеме верхняя граница обозначена вектором

$$\mathbf{U} = (U_1, U_2, \dots, U_n), \text{ а нижняя – вектором } \mathbf{L} = [L_1, L_2, \dots, L_n].$$

Используя преобразование (2), нижние границы α и β в значении $-0,3$, а верхние – в $0,3$, можно получить нижние и верхние границы изменения параметров функций принадлежности.

Если параметры выходят за границы требуемого диапазона, то принимают граничные значения. Сам спуск является итерационным процессом, в котором $\mathbf{P}$ изменяется следующим образом:

$$\mathbf{P} = \mathbf{P} - h \nabla E(\mathbf{P}), \quad (8)$$

где h – длина шага в направлении спуска, $\nabla E(\mathbf{P}) = (E'_{p_1}(\mathbf{P}), E'_{p_2}(\mathbf{P}), \dots, E'_{p_n}(\mathbf{P}))$ – градиент, элементы которого вычисляются согласно (5), (6) и (7). Длина шага в предлагаемом алгоритме выбирается методом дробления, т.е. итерации с преобразованиями (8) проводятся до тех пор, пока функция E убывает и не уменьшается точность классификации $acc(\mathbf{P})$, определяющая процент правильно классифицированных образцов обучающих данных. Если убывание прекращается или точность уменьшается, то шаг h уменьшается вдвое, вычисляются параметры $\mathbf{P}$ с новым шагом и снова проверяется условие. Дробление шага продолжается до тех пор, пока не будет выполнено условие. Итерационный процесс прекращается по достижению длины шага h минимального допустимого значения e , которое задается до выполнения работы алгоритма.

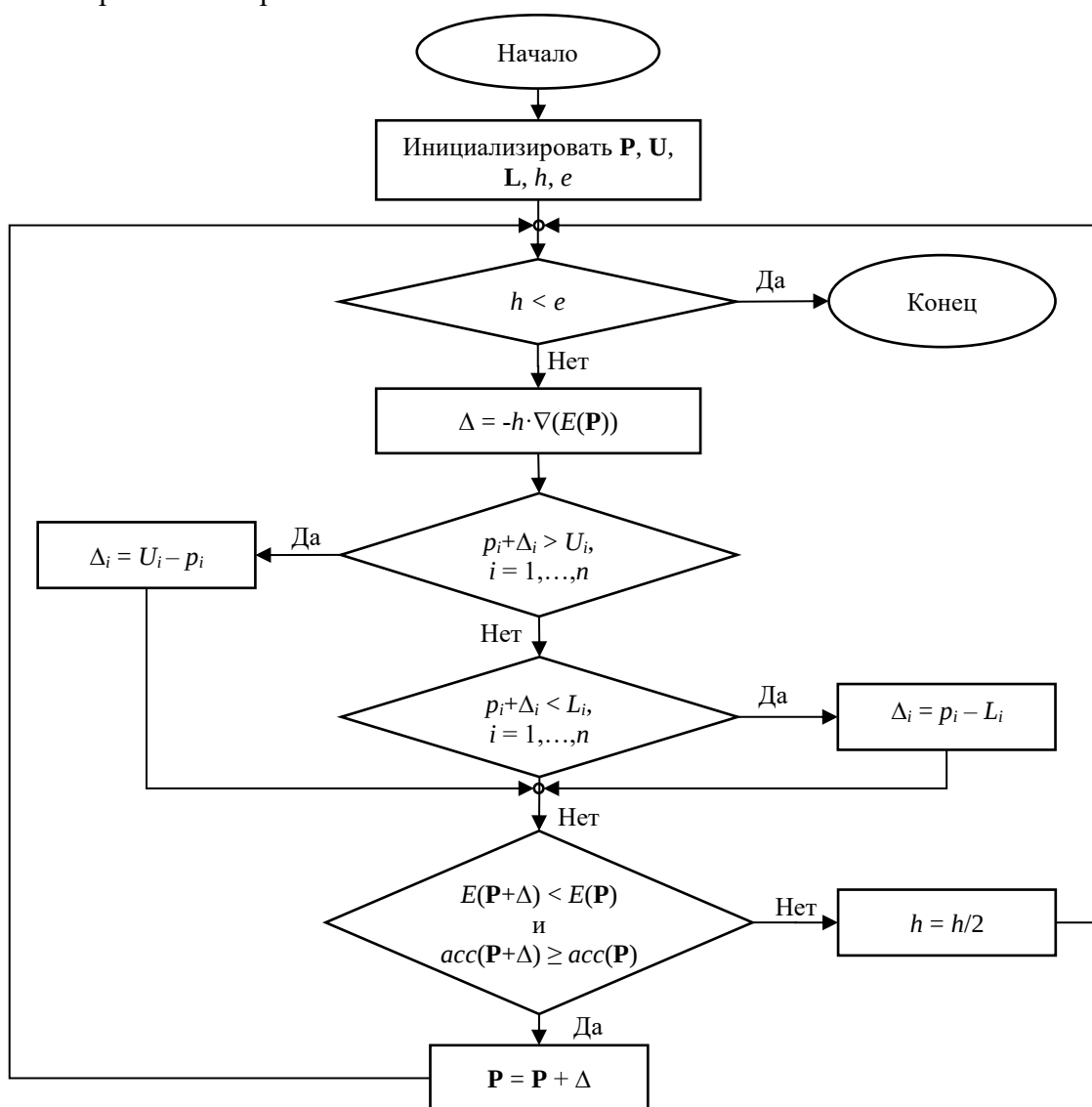


Рис. 2. Блок-схема алгоритма градиентного спуска

3. Эксперимент. Для оценки эффективности работы предлагаемого алгоритма настройки параметров, проведен эксперимент по построению нечетких классификаторов на общедоступных наборах. Настройка параметров градиентным спуском проводилась после

построения классификаторов алгоритмом смешанной многокритериальной оптимизации. Параметры алгоритмов указаны в таблице 1. Для улучшения интерпретируемости в алгоритме смешанной многокритериальной оптимизации генерировались правила, не содержащие больше трех лингвистических терминов, и пространство каждой переменной разбивалось на три нечетких множества. Нижние и верхние границы изменения параметров функций принадлежности L и U рассчитывались согласно (2), исходя из диапазона изменений $[-0,3, 0,3]$ бокового смещения α и смещения охвата β .

Таблица 1. Параметры алгоритмов построения

Алгоритм	Название параметра	Значение параметра
Смешанная многокритериальная оптимизация [26]	Размер популяции	100
	Число итераций	5000
	Вероятность удаления гнезда p_a	0,05
	Коэффициент прыжка полетов Леви γ	Равномерно уменьшался с 0,4 до 0,05
Градиентный спуск	Минимальная длина шага e	$1 \cdot 10^{-8}$
	Начальное значение шага h	0,1
	Диапазон изменения смещений α и β	$[-0,3, 0,3]$

3.1. Методика проведения эксперимента и результаты. Программная реализация осуществлялась на языке MATLAB в среде программирования MATLAB R2022b. Эксперимент проводился на персональном компьютере под управлением операционной системы Windows 10 с 8 Гб оперативной памяти и процессором Intel Core i7-12700.

Для экспериментов были использованы 38 наборов данных из репозитория KEEL (<http://keel.es>). Количество признаков в наборах варьируется от 2 до 90; количество образцов от 80 до 19020; количество классов от 2 до 15. Три набора являются синтетическими, т.е. специально разработанными для проверки эффективности классификаторов, остальные собраны из реальных проблемных областей.

Эксперимент проводился по схеме 10-кратной перекрестной проверки и соответствовал методике, приведенной в [29]. Для каждого из 10 разбиений набора данных запуск алгоритмов осуществлялся три раза. Таким образом, критерий эффективности оценивался средним значением по тридцати запускам алгоритма на каждом наборе данных. В качестве критерия эффективности применялся показатель точности – процент правильно классифицированных образцов. В таблице 2 приведены значения точности в ходе проведения эксперимента. В столбцах MI указана точность классификаторов, построенных метаэвристическим алгоритмом, а в столбцах GD указана точность классификаторов, дообученных градиентным спуском.

3.2. Обсуждение результатов и выводы. Для того, чтобы оценить значимость различий критериев эффективности при построении классификаторов, применяют статистический анализ, в частности, непараметрические тесты [30-32]. Для парных сравнений результатов эксперимента использовался тест знаковых рангов Уилкоксона, который основан на вычислении разности между показателями эффективности. Установлен уровень значимости $\alpha = 0,05$. Нулевая гипотеза H_0 говорит об отсутствии статистической разницы в результатах точности классификаторов от их дообучения алгоритмом градиентного спуска. Альтернативная гипотеза H_1 говорит об обратном. Результаты теста приведены в таблице 3. Суммарное количество отрицательных рангов оказалось больше положительных, а значение p -value оказалось меньше критического $\alpha = 0,05$. Таким образом, можно утверждать, что применение адаптированного алгоритма градиентного спуска для настройки параметров нечетких классификаторов, построенных алгоритмом смешанной многокритериальной

оптимизации «кукушкин поиск» с генетическим оператором скрещивания, показало статистически значимое увеличение точности классификации.

Таблица 2. Результаты точности классификации (%)

№	Набор данных	MI	GD	№	Набор данных	MI	GD
1	appendicitis	85,58	85,06	20	newthyroid	94,41	94,73
2	australian	85,60	86,09	21	page-blocks	94,72	94,85
3	balance	90,66	91,25	22	phoneme	79,21	79,35
4	banana	76,31	80,27	23	pima	75,92	76,23
5	bands	68,11	67,63	24	ring	92,69	92,80
6	cleveland	55,69	56,62	25	satimage	82,85	83,19
7	contraceptive	53,95	54,11	26	sonar	73,94	75,90
8	dermatology	92,55	94,03	27	spambase	89,91	89,99
9	ecoli	77,58	80,18	28	spectfheart	78,18	78,95
10	glass	64,44	65,03	29	tae	52,83	52,83
11	haberman	71,45	74,04	30	thyroid	94,30	94,30
12	hayes-roth	78,33	79,79	31	titanic	78,69	78,82
13	heart	82,72	83,33	32	twonorm	94,04	94,08
14	hepatitis	86,71	85,98	33	vehicle	66,20	67,66
15	ionosphere	91,37	91,28	34	vowel	42,56	43,94
16	iris	95,78	96,44	35	wdbc	96,08	96,31
17	magic	82,65	83,22	36	wine	94,53	96,24
18	mammographic	83,45	84,33	37	wisconsin	96,12	96,07
19	movement_libras	45,56	50,00	38	yeast	53,94	54,56

Таблица 3. Статистическое сравнение точности классификации

Сравнение алгоритмов	Сумма рангов		p-value	Гипотеза Н ₀ ($\alpha=0,05$)
	+	–		
MI и GD	75,5	590,5	<0,001	Отвергается

Для демонстрации свойства объяснимости нечетких классификаторов на рисунке 3 приведена база правил одного из вариантов построения для набора данных sonar.

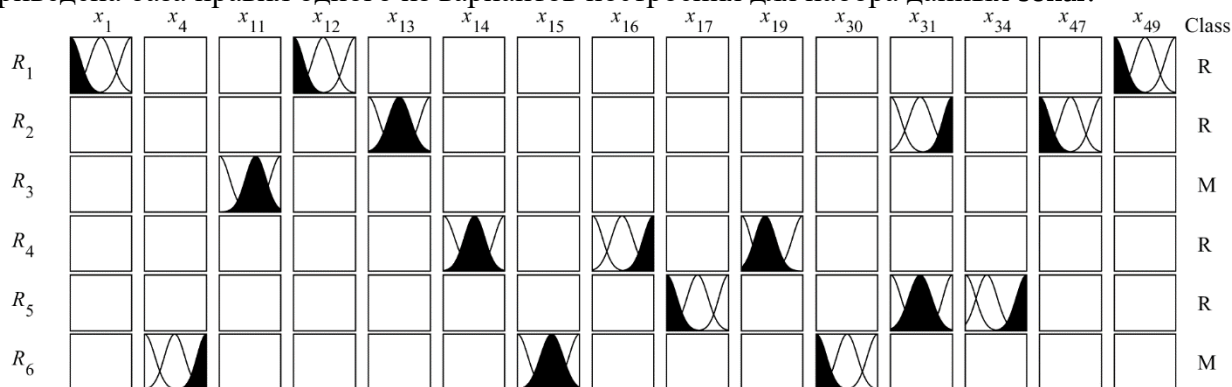


Рис. 3. Графическое представление базы правил нечеткого классификатора

Набор содержит 60 признаков, являющихся сигналами гидролокатора, и 208 образцов. Задачей классификации является определение по признакам гидролокатора, является обнаруженный объект камнем (R) или миной (M). Получены 6 правил ($R_1, R_2, \dots, R_6$). В конце каждого правила указано значение его класса (R или M), в столбцах размещены переменные. Как можно заметить, не все переменные задействованы в классификаторе. Приведены значения функций принадлежности для лингвистических терминов «Маленькое», «Среднее», «Большое». Функция принадлежности выбранного в правиле термина выделена заливкой. База правил на естественном языке представлена ниже.

- R_1 : ЕСЛИ x_1 Малое И x_{12} Малое И x_{49} Малое ТО объект является КАМНЕМ
 R_2 : ЕСЛИ x_{13} Среднее И x_{31} Большое И x_{47} Малое ТО объект является КАМНЕМ
 R_3 : ЕСЛИ x_{11} Среднее ТО объект является МИНОЙ
 R_4 : ЕСЛИ x_{14} Среднее И x_{16} Большое И x_{19} Среднее ТО объект является КАМНЕМ
 R_5 : ЕСЛИ x_{17} Малое И x_{31} Среднее И x_{34} Большое ТО объект является КАМНЕМ
 R_6 : ЕСЛИ x_4 Большое И x_{15} Среднее И x_{30} Малое ТО объект является МИНОЙ

Заключение. Для обеспечения доверия к системам искусственного интеллекта необходимо предоставлять доказательства или обоснования к результатам их решений. Такое требование делает актуальной разработку прогностических моделей на основе нечетких систем. Построение нечетких классификаторов сосредоточено на применении метаэвристических алгоритмов оптимизации, благодаря способности находить глобальный оптимум, работать в режиме «черного ящика» и обеспечивать поиск в дискретном и непрерывном пространстве. Предложенный адаптированный алгоритм градиентного спуска для настройки параметров функций принадлежности позволил статистически значимо улучшить точность классификации после построения метаэвристическим алгоритмом. Это осуществилось благодаря способности градиентного спуска точнее находить локальные оптимумы. Таким образом, применение комбинации метаэвристических и градиентных методов оптимизации обеспечивает увеличение точности прогнозов моделей на основе нечетких систем.

Благодарности. Автор выражает глубокую признательность Илье Александровичу Ходашинскому, профессору Томского государственного университета систем управления и радиоэлектроники, за обсуждение настоящего исследования и ряда ценных советов и замечаний. Исследование выполнено за счет гранта Российского научного фонда № 24-21-00168, <https://rscf.ru/project/24-21-00168/>.

Список источников

1. Wiese T. Predictive maintenance using Artificial Intelligence in critical infrastructure: a decision-making framework. International journal of engineering, business and management, 2024, vol. 8, no 4, pp. 1-4, DOI: 10.22161/ijebm.8.4.1.
2. Goodman B., Flaxman S. European union regulations on algorithmic decision making and a “right to explanation”. AI Magazine, 2017, vol. 38, no 3, pp. 50-57, DOI: 10.1609/aimag.v38i3.2741.
3. Zhu J., Liapis A., Risi S. et al. Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. 2018 IEEE Conference on Computational Intelligence and Games (CIG), 2018, pp. 1-8, DOI: 10.1109/CIG.2018.8490433.
4. Woodruff A., Fox S., Rousso-Schindler S. et al. A qualitative exploration of perceptions of algorithmic fairness. Proceedings Conference on human factors in computing systems, 2018, pp. 1-14, DOI: 10.1609/aimag.v38i3.2741.
5. Ходашинский И.А. Этические проблемы искусственного интеллекта / И.А. Ходашинский, Ю.А. Шурыгин, К.С. Сарин и др. // Современное образование: интеграция образования, науки, бизнеса и власти. Трансформация образования, науки и производства - основа технологического прорыва: материалы международной научно-методической конференции. В 2 ч., Томск, 26-27 января 2023 года. Том Часть 2. – Томск: Томский государственный университет систем управления и радиоэлектроники, 2023. – С. 125-129. – EDN XDAZY0.
6. Gunning D., Aha W. DARPA's explainable Artificial Intelligence (XAI) program. AI Magazine, 2019, vol. 40, pp. 44-58, DOI: 10.1609/aimag.v40i2.2850.
7. Ahmed I., Jeon G., Piccialli F. From Artificial Intelligence to explainable Artificial Intelligence in industry 4.0: a survey on what, how, and where. IEEE Transactions on industrial informatics, 2022, vol. 18, pp. 5031-5042, DOI: 10.1109/TII.2022.3146552.
8. Moral J.M.A., Castiello C., Magdalena L. et al. Explainable fuzzy systems. studies in computational intelligence. Springer: Cham, 2021, vol. 970, 232 p., DOI: 10.1007/978-3-030-71098-9.
9. Borisov V., Leemann T., Sesler K. et al. Deep neural networks and tabular data: a survey. IEEE Transactions on neural networks and learning systems, 2024, vol. 35, no. 6, pp. 7499-7519, DOI: 10.1109/TNNLS.2022.3229161.

10. Castelvechi D. Can we open the black box of AI? *Nature*, 2016, vol. 538, pp. 20-23, DOI:10.1038/538020a.
11. Lipton Z.C. The mythos of model interpretability. *Communications of the ACM*, 2016, vol. 61, pp. 36-43, DOI:10.1145/3233231.
12. Zadeh L.A. From computing with numbers to computing with words. From manipulation of measurements to manipulation of perception. *IEEE Transactions on circuits and systems I: fundamental theory and applications*, 1999, vol. 46, no 1, pp. 105-119, DOI: 10.1109/81.739259.
13. Zadeh L.A. Toward human-level machine intelligence: is it achievable? The need for a paradigm shift. *IEEE Computational Intelligence Magazine*, 2008, vol.3, no 3, pp. 11-22, DOI: 10.1109/MCI.2008.926583.
14. Casalino G., Castellano G., Castiello C. et al. Effect of fuzziness in fuzzy rule-based classifiers defined by strong fuzzy partitions and winner-takes-all inference. *Soft Computing*, 2022, vol. 26, pp. 6519-6527, DOI: 10.1007/s00500-022-07128-2.
15. Дуда Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт; Пер. с англ. Г. Г. Вайнштейна и А.М. Васьковского; Под ред. В.Л. Стефанюка. – Москва: Мир, 1976. – 511 с.
16. Herrera F. Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolution Intelligence*, 2008, vol. 1, no. 1, pp. 27-46, DOI: 10.1007/s12065-007-0001-5.
17. Jang R. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on systems, man and cybernetics*, 1993, vol.23, pp. 665-685, DOI: 10.1109/21.256541.
18. Cui Y., Wu D., Huang J. Optimize TSK fuzzy systems for classification problems: minibatch gradient descent with uniform regularization and batch normalization. *IEEE Transactions on fuzzy systems*, 2020, vol. 28, no. 12, pp. 3065-3075, DOI: 10.48550/arXiv.1908.00636.
19. Haque F., Reaz M.B.I., Chowdhury M.E.H. et al. Diabetic sensorimotor polyneuropathy severity classification using adaptive neuro fuzzy inference system. *IEEE Access*, 2021, vol. 9, pp. 7618-7631, DOI: 10.1109/ACCESS.2020.3048742.
20. Kanwal S., Jiriwibhakorn S. Advanced fault detection, classification, and localization in transmission lines: a comparative study of ANFIS, neural networks, and hybrid methods. *IEEE Access*, 2024, vol. 12, pp. 49017-49033, DOI: 10.1109/ACCESS.2024.3384761.
21. Бардамова М.Б. Формирование структуры нечёткого классификатора комбинацией алгоритма экстремумов классов и алгоритма «прыгающих лягушек» для несбалансированных данных с двумя классами / М.Б. Бардамова, И.А. Ходашинский // *Автометрия*, 2021. – Т. 57. – № 4. – С. 54-64. – DOI: 10.15372/AUT20210407. – EDN: BDRGMV.
22. Корышев Н.П. Алгоритм формирования базы правил нечёткого классификатора на основе алгоритма кластеризации K-средних и метаэвристического алгоритма «китов» / Н.П. Корышев, И.А. Ходашинский // *Доклады Томского государственного университета систем управления и радиоэлектроники*, 2021. – Т. 24. – № 1. – С. 42-47. – DOI: 10.21293/1818-0442-2021-24-1-42-47.
23. Остапенко Р.О. Построение нечёткого классификатора на основе метаэвристического алгоритма «саранчи» / Р.О. Остапенко, И.А. Ходашинский, Ю.А. Шурыгин // *Научно-техническая информация. Серия 2: Информационные процессы и системы*, 2023. – № 11. – С. 26-40. – DOI: 10.21293/1818-0442-2022-25-2-31-36.
24. Black box optimization, machine learning, and no-free lunch theorems, Eds. P.M. Pardalos, V. Rasskazova, M.N. Vrahatis. Cham: Springer, Springer optimization and its applications, 2021, vol. 170, 388 p., DOI: 10.1007/978-3-030-66515-9.
25. Ishibuchi H., Narashima T., Morisawa T. Voting in fuzzy rule-based systems for pattern classification problem. *Fuzzy sets and systems*, 1999, vol. 103, no 1, pp. 223-238, DOI: 10.1016/S0165-0114(98)00223-1.
26. Сарин К.С. Гибридный алгоритм смешанной многокритериальной оптимизации «кукушкин поиск» с генетическим оператором скрещивания / К.С. Сарин // *Искусственный интеллект и принятие решений*, 2024. – № 2. – С. 87-105. – DOI: 10.14357/20718594240207. – EDN: VQYBOD
27. Herrera F., Martinez-Lopez L. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on fuzzy systems*, 2000, vol. 8, iss. 6, pp. 746-752, DOI: 10.1109/91.890332.
28. Ishibuchi H., Nakashima T., Nii M. Classification and modeling with linguistic information granules. Berlin: Springer, 2005, 308 p., DOI: 10.1007/b138232.
29. Alcalá R., Nojima Y., Herrera F. et al. Multiobjective genetic fuzzy rule selection of single granularity-based fuzzy classification rules and its interaction with the lateral tuning of membership functions. *Soft Computing*, 2011, vol. 15, pp. 2303-2318, DOI: 10.1007/s00500-010-0671-2.
30. Demsar J. Statistical comparisons of classifiers over multiple data sets. *Journal of machine learning research*, 2006, vol. 7, pp. 1-30.
31. Garsia S., Herrera F. An extension on “statistical comparisons of classifiers over multiple data sets” for all pairwise comparisons. *Journal of machine learning research*, 2008, vol. 9, pp. 2677-2694.

32. Garsia S., Fernandez A., Luengo J. et al. Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. Information Sciences, 2010, vol. 180, pp. 2044-2064, DOI:10.1016/j.ins.2009.12.010.

Сарин Константин Сергеевич. Доцент, кандидат технических наук, доцент кафедры компьютерных систем в управлении и проектировании (КСУП) Томского государственного университета систем управления и радиоэлектроники (ТУСУР). Область научных исследований связана с интеллектуальным анализом данных, нечеткими системами и методами оптимизации. AuthorID: 947352, SPIN: 4838-1814, ORCID: 0000-0003-3264-7768, sarin.konstantin@mail.ru, Россия, г. Томск, ул. Вершинина, 74.

UDC 004.85

DOI:10.25729/ESI.2025.38.2.002

Adapted gradient descent algorithm for tuning parameters of fuzzy classifier

Konstantin S. Sarin

Tomsk state university of control systems and radioelectronics,
Russia, Tomsk, sarin.konstantin@mail.ru

Abstract. The use of artificial intelligence systems for critical areas of human life requires trust in the result obtained by the system. An explanation of the obtained solution is important for ensuring trust. Fuzzy systems have the property of explainability due to the presence of a base of production rules in natural language. This work is devoted to the development of a gradient descent algorithm for adjusting the parameters of fuzzy classifiers. Experiments on 38 data sets showed that the use of the developed algorithm for classifiers built by a metaheuristic algorithm statistically significantly increases the accuracy of classification.

Keywords: fuzzy systems, classification, data analysis, optimization algorithms

Acknowledgements: The author expresses deep gratitude to Ilya Aleksandrovich Hodashinsky, professor of Tomsk State University of Control Systems and Radioelectronics, for the discussion of this study and a number of valuable advice and comments. The study was supported by the grant of the Russian Science Foundation No. 24-21-00168, <https://rscf.ru/project/24-21-00168/>.

References

1. Wiese T. Predictive maintenance using Artificial Intelligence in critical infrastructure: a decision-making framework. International journal of engineering, business and management, 2024, vol. 8, no 4, pp. 1-4, DOI: 10.22161/ijebm.8.4.1.
2. Goodman B., Flaxman S. European union regulations on algorithmic decision making and a “right to explanation”. AI Magazine, 2017, vol. 38, no 3, pp. 50-57, DOI: 10.1609/aimag.v38i3.2741.
3. Zhu J., Liapis A., Risi S. et al. Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. 2018 IEEE Conference on Computational Intelligence and Games (CIG), 2018, pp. 1-8, DOI: 10.1109/CIG.2018.8490433.
4. Woodruff A., Fox S., Rousso-Schindler S. et al. A qualitative exploration of perceptions of algorithmic fairness. Proceedings Conference on human factors in computing systems, 2018, pp. 1-14, DOI: 10.1609/aimag.v38i3.2741.
5. Hodashinsky I.A., Shurygin Yu.A., Sarin K.S. et al. Eticheskie problemy iskusstvennogo intellekta [Ethical Issues of Artificial Intelligence]. Sovremennoe obrazovanie: integraciya obrazovaniya, nauki, biznesa i vlasti. Transformaciya obrazovaniya, nauki i proizvodstva - osnova tekhnologicheskogo proryva: materialy mezhdunarodnoj nauchno-metodicheskoy konferencii [Modern education: integration of education, science, business and government. Transformation of education, science and production - the basis of technological breakthrough: materials of the international scientific and methodological conference]. Tomsk, 26-27 yanvarya 2023 goda, Tomsk: Tomskiy gosudarstvennyy universitet sistem upravleniya i radioelektroniki [Tomsk, 26-27 January 2023, Tomsk: Tomsk state university of control systems and radioelectronics], 2023, vol 2, pp. 125-129, EDN XDAZYO.
6. Gunning D., Aha W. DARPA's explainable Artificial Intelligence (XAI) program. AI Magazine, 2019, vol. 40, pp. 44-58, DOI: 10.1609/aimag.v40i2.2850.

7. Ahmed I., Jeon G., Piccialli F. From Artificial Intelligence to explainable Artificial Intelligence in industry 4.0: a survey on what, how, and where. *IEEE Transactions on industrial informatics*, 2022, vol. 18, pp. 5031-5042, DOI: 10.1109/TII.2022.3146552.
8. Moral J.M.A., Castiello C., Magdalena L. et al. *Explainable fuzzy systems. studies in computational intelligence*. Springer: Cham, 2021, vol. 970, 232 p., DOI: 10.1007/978-3-030-71098-9.
9. Borisov V., Leemann T., Sesler K. et al. Deep neural networks and tabular data: a survey. *IEEE Transactions on neural networks and learning systems*, 2024, vol. 35, no. 6, pp. 7499-7519, DOI: 10.1109/TNNLS.2022.3229161.
10. Castelvechi D. Can we open the black box of AI? *Nature*, 2016, vol. 538, pp. 20-23, DOI:10.1038/538020a.
11. Lipton Z.C. The mythos of model interpretability. *Communications of the ACM*, 2016, vol. 61, pp. 36-43, DOI:10.1145/3233231.
12. Zadeh L.A. From computing with numbers to computing with words. From manipulation of measurements to manipulation of perception. *IEEE Transactions on circuits and systems I: fundamental theory and applications*, 1999, vol. 46, no 1, pp. 105-119, DOI: 10.1109/81.739259.
13. Zadeh L.A. Toward human-level machine intelligence: is it achievable? The need for a paradigm shift. *IEEE Computational Intelligence Magazine*, 2008, vol.3, no 3, pp. 11-22, DOI: 10.1109/MCI.2008.926583.
14. Casalino G., Castellano G., Castiello C. et al. Effect of fuzziness in fuzzy rule-based classifiers defined by strong fuzzy partitions and winner-takes-all inference. *Soft Computing*, 2022, vol. 26, pp. 6519-6527, DOI: 10.1007/s00500-022-07128-2.
15. Duda R., Hart P. *Распознавание образов и анализ стсен [Pattern classification and scene analysis]*. Москва: Mir [Moscow: Mir], 1976, 511 p.
16. Herrera F. Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolution Intelligence*, 2008, vol. 1, no. 1, pp. 27-46, DOI: 10.1007/s12065-007-0001-5.
17. Jang R. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on systems, man and cybernetics*, 1993, vol.23, pp. 665-685, DOI: 10.1109/21.256541.
18. Cui Y., Wu D., Huang J. Optimize TSK fuzzy systems for classification problems: minibatch gradient descent with uniform regularization and batch normalization. *IEEE Transactions on fuzzy systems*, 2020, vol. 28, no. 12, pp. 3065-3075, DOI: 10.48550/arXiv.1908.00636.
19. Haque F., Reaz M.B.I., Chowdhury M.E.H. et al. Diabetic sensorimotor polyneuropathy severity classification using adaptive neuro fuzzy inference system. *IEEE Access*, 2021, vol. 9, pp. 7618-7631, DOI: 10.1109/ACCESS.2020.3048742.
20. Kanwal S., Jiriwibhakorn S. Advanced fault detection, classification, and localization in transmission lines: a comparative study of ANFIS, neural networks, and hybrid methods. *IEEE Access*, 2024, vol. 12, pp. 49017-49033, DOI: 10.1109/ACCESS.2024.3384761.
21. Bardamova M.B., Hodashinsky I.A. Formirovaniye struktury nechotkogo klassifikatora kombinatsiyey algoritma ekstremumov klassov i algoritma "prygayushchikh lyagushek" dlya nesbalansirovannykh dannyykh s dvumya klassami [Formation of fuzzy classifier structure by a combination of the class extremum algorithm and the shuffled frog leaping algorithm for imbalanced data with two classes]. *Avtometriya [Avtometriya]*, 2021, vol. 57, no. 4, pp. 378-387, DOI: 10.15372/AUT20210407, EDN: BDRGMV.
22. Koryshev N.P., Hodashinsky I.A. Algoritm formirovaniya bazy pravil nechotkogo klassifikatora na osnove algoritma klasterizatsii K-srednih i metaevristicheskogo algoritma "kitov" [Algorithm to forming a rule base for a fuzzy classifier de-signed on the basis of the K-means clustering algorithm and the whale optimization algorithm]. *Doklady TUSUR [Reports of TUSUR]*, 2021, vol. 24, no 1, pp. 42-47, DOI: 10.21293/1818-0442-2021-24-1-42-47.
23. Ostapenko R.O., Hodashinsky I.A., Shurygin Yu.A. Postroyeniye nechotkogo klassifikatora na osnove metaevristicheskogo algoritma "saranchi" [Using the grasshopper optimization algorithm for fuzzy classifier design]. *Nauchno-tekhnicheskaya informatsiya. Seriya 2: Informatsionnyye protsessy i sistemy [Scientific and technical information. Series 2: Information processes and systems]*, 2023, no. 11, pp. 26-40, DOI: 10.21293/1818-0442-2022-25-2-31-36.
24. Black box optimization, machine learning, and no-free lunch theorems, Eds. P.M. Pardalos, V. Rasskazova, M.N. Vrahatis. Cham: Springer, Springer optimization and its applications, 2021, vol. 170, 388 p., DOI: 10.1007/978-3-030-66515-9.
25. Ishibuchi H., Narashima T., Morisawa T. Voting in fuzzy rule-based systems for pattern classification problem. *Fuzzy sets and systems*, 1999, vol. 103, no 1, pp. 223-238, DOI: 10.1016/S0165-0114(98)00223-1.
26. Sarin K.S. Gibridnyj algoritm smeshannoj mnogokriterial'noj optimizatsii "kukushkin poisk" s geneticheskim operatorom skreshchivaniya [Hybrid algorithm of mixed-integer multi-objective optimization "cuckoo search" with genetic crossover operator]. *Iskusstvennyj intellekt i prinyatie reshenij [Artificial intelligence and decision making]*, 2024, no 2, pp. 87-105, DOI: 10.14357/20718594240207, EDN: VQYBOD.

27. Herrera F., Martinez-Lopez L. A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Transactions on fuzzy systems, 2000, vol. 8, iss. 6, pp. 746-752, DOI: 10.1109/91.890332.
28. Ishibuchi H., Nakashima T., Nii M. Classification and modeling with linguistic information granules. Berlin: Springer, 2005, 308 p., DOI: 10.1007/b138232.
29. Alcalá R., Nojima Y., Herrera F. et al. Multiobjective genetic fuzzy rule selection of single granularity-based fuzzy classification rules and its interaction with the lateral tuning of membership functions. Soft Computing, 2011, vol. 15, pp. 2303-2318, DOI: 10.1007/s00500-010-0671-2.
30. Demsar J. Statistical comparisons of classifiers over multiple data sets. Journal of machine learning research, 2006, vol. 7, pp. 1-30.
31. Garsia S., Herrera F. An extension on “statistical comparisons of classifiers over multiple data sets” for all pairwise comparisons. Journal of machine learning research, 2008, vol. 9, pp. 2677-2694.
32. Garsia S., Fernandez A., Luengo J. et al. Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. Information Sciences, 2010, vol. 180, pp. 2044-2064, DOI:10.1016/j.ins.2009.12.010.

Sarin Konstantin Sergeevich. Associate professor, candidate of technical sciences, associate professor of the department of computer systems in control and design (KSUP) of Tomsk state university of control systems and radioelectronics (TUSUR). The area of scientific research is related to data mining, fuzzy systems and optimization methods. AuthorID: 947352, SPIN: 4838-1814, ORCID: 0000-0003-3264-7768, sarin.konstantin@mail.ru, Russia, Tomsk, Vershinina St., 74.

Статья поступила в редакцию 03.03.2025; одобрена после рецензирования 26.03.2025; принята к публикации 10.04.2025.

The article was submitted 03/03/2025; approved after reviewing 03/26/2025; accepted for publication 04/10/2025.