



**Российский университет
дружбы народов (РУДН)**

ул. Минлухо-Маклая, д. 6, Москва, Россия, 117198
ОГРН 1027739189323; ОКПО 02066463; ИНН 7728073720

Телефон: +7495 434 53 00, факс: +7495 433 15 11
www.rudn.ru; rudn@rudn.ru

14 ноября 2022
№ 05-07 / 6607

УТВЕРЖДАЮ
Первый проректор -
проректор по научной работе,
доктор медицинских наук, профессор,
член-корреспондент РАН

Костин Андрей Александрович

«14» ноября 2022 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ
федерального государственного автономного образовательного
учреждения высшего образования
«Российский университет дружбы народов» (РУДН)
на диссертацию Девяткина Дмитрия Алексеевича на тему
«Построение ансамблей деревьев решений с использованием линейных и
нелинейных разделителей»,
представленную к защите в диссертационном совете Д 24.1.120.01 на базе
Федерального государственного бюджетного учреждения науки
Института системного программирования им. В.П. Иванникова
Российской академии наук (ИСП РАН)
на соискание ученой степени кандидата физико-математических наук
по специальности 2.3.5 – Математическое и программное обеспечение
вычислительных систем, комплексов и компьютерных сетей

Актуальность темы диссертационной работы

Диссертационная работа Девяткина Д.А. направлена на решение проблемы повышения качества классификации на основе лесов деревьев решений путем создания вычислительно-эффективного метода построения случайных ансамблей деревьев решений с применением линейных и нелинейных разделителей. В ряде задач интеллектуального анализа данных

существенное значение имеет дискретная природа составляющих частей анализируемых объектов. Высокой точности решения в таких случаях можно достигнуть при применении моделей машинного обучения с дискретными зависимыми переменными. В основе таких моделей могут лежать деревья решений, однако они имеют ограниченную выразительную способность при фиксированной высоте. Одним из подходов к решению проблемы является обучение деревьев решений с многомерными разделителями в узлах, однако подобные деревья решений имеют низкую обобщающую способность, поэтому они применимы только в составе рандомизированных композиций, при условии использования методов редукции сложности, позволяющих найти баланс между сложностью получаемых алгоритмов и их обобщающей способностью. Кроме того, большинство подходов к построению таких деревьев имеют низкую вычислительную эффективность и большое количество гиперпараметров. Поэтому задача создания и развития методов построения случайных ансамблей деревьев решений с применением линейных и нелинейных разделителей является актуальной.

Научная новизна полученных результатов

Автором диссертационного исследования разработан оригинальный вычислительно-эффективный метод построения узлов деревьев решений с применением линейных и нелинейных разделителей, при обучении которых совместно оптимизируется отступ между разделяемыми объектами и произвольный критерий однородности. Этот метод применен для обучения деревьев решений в составе случайных лесов. Выполнено развитие теоретических подходов к подбору методов редукции сложности случайных ансамблей деревьев решений: теоретически обоснована связь между равномерной стабильностью алгоритмов обучения и формируемой структурой деревьев решений, предложена оценка обобщающей способности случайных ансамблей деревьев решений, учитывающая основные гиперпараметры алгоритмов их построения. Предложена архитектура программного обеспечения глобально распределенного обучения случайных лесов деревьев решений с линейными и нелинейными разделителями.

Значимость полученных результатов для науки и практики

Представленные результаты диссертационной работы состоят в создании новых методов, разработке архитектуры и программных средств интеллектуального анализа данных и в развитии формальных подходов к исследованию обобщающей способности случайных ансамблей деревьев

решений. Эти подходы могут быть применены для создания или подбора методов регуляризации случайных ансамблей деревьев решений.

Разработанные методы, архитектура и программные средства могут применяться на практике при создании систем интеллектуального анализа данных, элементы которых имеют дискретную природу. Результаты диссертационной работы рекомендуется применять для автоматизации различных операций в промышленном производстве, сельском хозяйстве, энергетике, на транспорте.

Результаты работы соответствуют приоритетному направлению развития науки, технологий и техники из Стратегии научно-технологического развития Российской Федерации (утверждена Указом Президента Российской Федерации от 01.12.2016 №642 «О Стратегии научно-технологического развития Российской Федерации») «Переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, новым материалам и способам конструирования, создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта»; и могут внести вклад в развитие критических технологий Российской Федерации: пункт «8. Нано-, био-, информационные, когнитивные технологии».

Достоверность и обоснованность научных результатов

Обоснованность положений диссертации подтверждается глубокой проработкой научно-технических источников, строгими математическими формулировками положений, строгим математическим доказательством утверждений, лемм и теорем, корректностью проведенных экспериментальных исследований.

Достоверность результатов обеспечивается обоснованностью формальных доказательств. Достоверность обеспечивается апробацией на научных семинарах и международной конференции, успешным применением результатов при выполнении нескольких НИОКР, а также научными публикациями в рецензируемых журналах, шесть из которых опубликованы в журналах из перечня ВАК или приравненных к ним (Web of Science/Scopus, Q1/Q2).

Замечания по диссертационной работе

1. В главах 1 и 2 широко используется понятие «отступ анализируемых объектов». Необходимо дать определение этому понятию.
2. В главе 3 отсутствует анализ влияния выбранных значений гиперпараметров обучения лесов деревьев решений на их

- обобщающую способность с применением формального подхода, предложенного в главе 2.
3. Было бы полезно привести в главе 3 оценки качества работы оригинального случайного леса деревьев решений, усиленного с помощью подхода, выбранного в главе 2.
 4. Глава 3 содержит как описание архитектуры и реализации программы для обучения случайных лесов деревьев решений, так и результаты экспериментальных исследований. Для повышения связности материала в главах диссертации стоило бы вынести результаты экспериментальных исследований в отдельную главу.
 5. Текст работы содержит небольшое число опечаток и ошибок оформления. В обзорной части (глава 1) на стр. 27 и стр. 28 пропущены две ссылки на формулы.

Заключение

Отмеченные недостатки не влияют на общую положительную оценку диссертационной работы Девяткина Дмитрия Алексеевича. Диссертация «Построение ансамблей деревьев решений с использованием линейных и нелинейных разделителей» является научно-квалификационной работой, в которой содержится решение научной задачи повышения качества классификации методами на основе лесов деревьев решений, имеющей значение для отрасли цифрового развития, а также специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Работа выполнена на высоком научном уровне, имеет теоретическую и практическую значимость, изложена ясным и грамотным языком. Содержание автореферата соответствует основным положениям диссертации и корректно их отражает. Научная новизна и оригинальность результатов диссертационного исследования подтверждена публикациями в периодических изданиях, рекомендованных ВАК РФ и приравненных к ним. Публикации по теме диссертации в полной мере освещают ее основные положения. Полученные автором результаты апробированы на международной конференции и научных семинарах, опубликованы в 3 научных изданиях, входящих в базы Web of Science/Scopus (Q1/Q2), и в 3 журналах, рекомендованных ВАК, зарегистрированы в ФИПС в форме программы для ЭВМ.

На основании изложенного считаем, что диссертация Девяткина Дмитрия Алексеевича «Построение ансамблей деревьев решений с использованием линейных и нелинейных разделителей» соответствует

критериям, которые установлены пп. 9-14 Положения о присуждении ученых степеней (утв. Постановлением Правительства РФ от 24.09.2013 № 842, в редакции от 26.09.2022), предъявляемым в отношении диссертаций на соискание ученой степени кандидата физико-математических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей, а ее автор Девяткин Дмитрий Алексеевич заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Диссертация и автореферат заслушаны и обсуждены на заседании кафедры прикладной информатики и теории вероятностей федерального государственного автономного образовательного учреждения высшего образования «Российский университет дружбы народов» 08.11.2022 (состав кафедры – 46, присутствовали – 35), протокол № 0200-19-04/04.

08.11.2022

Доцент кафедры прикладной информатики
и теории вероятностей РУДН,
кандидат физико-математических наук
(специальность 05.13.17 – Теоретические
основы информатики), доцент
Кочеткова Ирина Андреевна

Федеральное государственное автономное образовательное учреждение
высшего образования «Российский университет дружбы народов» (ФГАОУ
ВО РУДН)

Адрес: 117198, г. Москва, ул. Миклухо-Маклая, д. 6

Тел.: +7 (495) 434-70-27

Сайт: <http://www.rudn.ru/>

E-mail: rector@rudn.ru