

ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

о диссертационной работе Ишкиной Шауры Хабировны
«Комбинаторные оценки переобучения пороговых решающих правил»,
представленной на соискание ученой степени кандидата физико-математических наук
по специальности 1.2.1 — «Искусственный интеллект и машинное обучение»

Целью диссертационной работы Ш. Х. Ишкиной является получение оценок обобщающей способности семейства одномерных пороговых решающих правил и их применение в практических задачах классификации. Работа выполнена в рамках комбинаторной теории надёжности обучения по прецедентам, которая позволяет получать точные (достигаемые) оценки вероятности переобучения и полного скользящего контроля. При этом оставался открытым вопрос, возможно ли использовать комбинаторные оценки для повышения качества логических методов классификации, использующих пороговые правила для бинаризации числовых признаков.

Актуальность работы связана со значительным расширением областей приложения машинного обучения и потребностью решать прикладные задачи с малым объёмом данных и высоким риском переобучения. Оценки обобщающей способности, получаемые в рамках статистической теории обучения, носят асимптотический характер и сильно завышены. Первые такие оценки были получены В. Н. Вапником и А. Я. Червоненкисом в конце 60-х годов. Несмотря на значительный прогресс в последующие годы, статистическая теория до сих пор не выработала оценок, адекватных в условиях малых данных. Комбинаторная теория переобучения обходится без асимптотических приближений, что позволяет вычислять оценки обобщающей способности по выборке. При некоторых условиях такие оценки являются точными, однако эти условия слишком специфичны и на практике редко выполняются. Например, оценки И. С. Гуза для одномерных пороговых решающих правил справедливы лишь при условии, что все значения признака на классифицируемых объектах попарно различны.

Ш. Х. Ишкиной удалось снять это ограничение, и для произвольного распределения значений одномерного признака в задаче бинарной классификации получить достигаемые оценки обобщающей способности одномерных пороговых решающих правил. Алгоритм вычисления этих оценок имеет полиномиальную сложность $O(L^5)$ по объёму выборки L , что делает их расчет на компьютере практически невозможным при объёмах более нескольких сотен объектов. Для решения этой проблемы в диссертационной работе предложена суррогатная модель, которая позволяет аппроксимировать оценки с высокой точностью ($\text{MAPE} = 2.8\%$), снизив сложность вычислений до $O(L^2)$. Такой подход в рамках комбинаторной теории реализован впервые.

Анализ значимости признаков в суррогатной модели позволил выявить признаки семейства классификаторов, которые можно использовать при построении новых оценок обобщающей способности для других семейств.

Практическая значимость данного диссертационного исследования заключается в том, что одномерные пороговые правила встречаются во многих алгоритмах машинного обучения

— решающих деревьях, решающих списках, голосовании элементарных конъюнкций, нейронных сетях с пороговыми функциями активации, а также при бинаризации данных. Применение оценок обобщающей способности позволяет строить алгоритмы классификации с контролируемым переобучением. Пример реализации приведен в диссертационной работе Ш. Х. Ишкиной, где показано, что использование комбинаторных оценок при выборе атрибута для разделения узла решающего дерева повышает его обобщающую способность.

Ш. Х. Ишкиной получен патент РФ на изобретение, где решающее дерево используется при построении программы трассерных (маркерных) исследований и позволяет уточнить список пар скважин с обнаруженным взаимовлиянием, что повышает эффективность разработки нефтегазового месторождения. Результаты диссертационной работы нашли свое применение при выполнении научно-исследовательских и инженерно-аналитических работ и учтены при разработке методических указаний в ООО «РН-БашНИПИнефть».

В ходе работы над диссертацией Ш. Х. Ишкина проявила себя как квалифицированный исследователь, самостоятельно ставила научные задачи, находила обоснованные способы их решения, проводила вычислительные эксперименты, оформляла результаты исследований в научных публикациях. Данное исследование вносит важный вклад в развитие комбинаторной теории переобучения и методологии планирования, проведения и интерпретации трассерных (маркерных) исследований.

Представленные в диссертации результаты получены автором самостоятельно, являются оригинальными и имеют как теоретическое, так и прикладное значение. Основные результаты опубликованы в 16 работах, из которых 8 работ представлены в рецензируемых изданиях из перечня ВАК и приравненных к ним международных наукометрических базах.

Работа соответствует всем требованиям, предъявляемым к диссертациям по специальности 1.2.1 — «Искусственный интеллект и машинное обучение». Диссертант Ишкина Ш. Х. является квалифицированным специалистом по данной специальности и заслуживает присвоения ей ученой степени кандидата физико-математических наук.

10 декабря 2025 г.

Заведующий кафедрой
«Математические методы прогнозирования»
факультета ВМК Московского государственного
университета имени М. В. Ломоносова

К. В. Воронцов

Подпись удостоверяю
декан факультета ВМК МГУ
академик РАН И. А. Соколов

