

УТВЕРЖДАЮ



Директор ИММ УрО РАН,
д.ф.-м.н., проф. РАН

Н.Ю. Лукоянов

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

о диссертационной работе Прокофьева Петра Александровича «КОРРЕКТНОЕ РАСПОЗНАВАНИЕ ПО ПРЕЦЕДЕНТАМ: ПОСТРОЕНИЕ ЛОГИЧЕСКИХ КОРРЕКТОРОВ ОБЩЕГО ВИДА И ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ», представленной на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 - «Теоретические основы информатики»

Диссертация П.А. Прокофьева посвящена решению серии актуальных дискретных задач, индуцированных проблемой обучения распознаванию в классе логических решающих правил. Подход к проблеме распознавания, в рамках которого при обучении классификатора используется индуктивный вывод логических закономерностей таких, как конъюнктивные правила, исторически является, по-видимому, одним из первых и наиболее глубоко исследованных в современной теории обучения. Моделям таких классификаторов посвящено большое число работ как отечественных (Ю.И. Журавлева, М.М. Бонгарда, М.Н. Вайнцвайга и др.), так и зарубежных исследователей. Тем не менее, ряд вопросов, возникающих на пути создания алгоритмов обучения, обладающих высокой эффективностью, обобщающей способностью и устойчивостью к случайным флуктуациям материала обучения, по-прежнему остаются открытыми. С математической точки зрения изучаемые в работе задачи тесно связаны с классической задачей дуализации булевой функции, привлекающих внимание исследователей в области дискретной математики долгие годы. Одна из причин столь пристального интереса к данной задаче состоит в том, что статус ее вычислительной сложности до сих пор не выяснен, и, следовательно, построение новых эффективных точных и приближенных алгоритмов ее решения приближает нас к ответу на вопрос о взаимодействии классов P и NP.

Таким образом, актуальность данного исследования, в рамках которого разработана серия новых логических классификаторов, позволяющих в определенной степени улучшить качество решения ряда

задач распознавания, а также усовершенствованы алгоритмы обработки дискретных данных для снижения трудоемкости обучения как ранее построенных, так и новых логических классификаторов, **не вызывает сомнений.**

Диссертационная работа состоит из введения, трех глав, заключения и списка литературы. Во введении автор обосновывает актуальность темы, новизну полученных результатов и приводит постановки основных задач.

Первая глава начинается с обзора распространенных подходов к построению логических классификаторов: логического, оптимизационного, алгебраического и алгебро-логического. Последний подход развивается в работах Е.В. Дюковой и ее учеников и фактически объединяет идеи логического и алгебраического подходов. Вводятся понятия элементарного классификатора (конъюнкции на множестве допустимых целочисленных описаний объектов) и корректного набора элементарных классификаторов (базовое понятие для алгебро-логического синтеза). Коррекция элементарных классификаторов осуществляется булевыми функциями.

Автор обобщает понятие корректного набора элементарных классификаторов и тем самым расширяет модель логических корректоров – алгоритмов голосования по корректным наборам элементарных классификаторов. Построена модель POLAR, в которой для коррекции используются поляризуемые булевы функции (ранее либо применялись только монотонные булевы функции, либо на корректирующие булевы функции не накладывалось никаких ограничений). Этап обучения классификатора POLAR сведен к поиску специальных покрытий булевой матрицы.

Вторая глава посвящена методам повышения скорости обучения и улучшения обобщающей способности логических корректоров. Предложено голосующие семейства наборов элементарных классификаторов логического корректора формировать методом бустинга (в более ранних работах применялся генетический алгоритм) и предварительно отбирать наиболее информативные элементарные классификаторы для каждого класса задачи – строить локальные базисы классов (ранее фактически использовались локальные базисы, состоящие из всевозможных элементарных классификаторов с одним признаком). Эффективность указанных приемов, носящих довольно универсальный характер, проиллюстрирована на логическом корректоре POLAR.

Глава содержит описание численного сравнения предложенных алгоритмов с известными аналогами по качеству классификации. При проведении эксперимента автором использованы методики скользящего

экзамена по фиксированному множеству сложных задач обучения (взятых из репозитория UCI и др. известных источников).

В третьей главе решается задача дуализации булевой матрицы (перечисление неприводимых покрытий булевой матрицы). Эта задача возникает как при обучении классических логических классификаторов, так и при построении логических корректоров. Полученные в данной главе результаты особенно актуальными, в связи с возможностью применения алгоритмов дуализации во многих других областях дискретного анализа данных, отличных от распознавания образов.

Традиционно эффективность перечислительных алгоритмов оценивается сложностью получения очередного решения в худшем случае. Существует два принципиально различных способа перечисления решений: с запоминанием решений, найденных на предыдущих шагах (инкрементальный подход) и без запоминания. Построению инкрементальных алгоритмов дуализации посвящено достаточно много работ (Л.Г. Хачиян, В.А. Гурвич, Е. Boros, К. Elbassioni и др.). Перечислять неприводимые покрытия, не запоминая ранее найденные, позволяют эффективные «в среднем» асимптотически оптимальные алгоритмы дуализации, предложенные Е.В. Дюковой. Автору удалось снизить трудоемкость ранее построенных асимптотически оптимальных алгоритмов, тем самым опередив алгоритмы «без памяти», занимающие до настоящего времени лидирующие позиции.

В целом диссертационная работа П.А. Прокофьева содержит решение нескольких актуальных задач, связанных с разработкой новых алгоритмов обучения логических классификаторов, алгоритмов дуализации и обоснованием оценок их вычислительной сложности, что вполне соответствует паспорту специальности 05.13.17.

Основные результаты диссертации являются новыми, имеют теоретическое и практическое значение. Они полностью опубликованы в 15 работах, в том числе, в 5 статьях из списка изданий, рекомендованных ВАК, докладывались на международных и всероссийских конференциях. Автореферат адекватно отражает содержание диссертации.

Принципиальных замечаний по содержанию работы и изложению материала нет.

Представленная диссертация является научной квалификационной работой, основные результаты которой можно рассматривать как решение актуальных задач в области алгоритмического обучения и анализа данных. Разработанные автором алгоритмы рекомендуются к использованию в Институте математики и механики им. Н.Н. Красовского УрО РАН, Институте

математики им. С.Л. Соболева СО РАН, Уральском федеральном университете и других организациях, а также к включению в состав соответствующих специальных курсов для студентов и магистрантов университетов, обучающихся по специальностям прикладная математика и теоретическая информатика.

Работа удовлетворяет требованиям, предъявляемым ВАК к кандидатским диссертациям, а ее автор, Петр Александрович Прокофьев, заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.17 – теоретические основы информатики.

Отзыв на диссертацию П.А. Прокофьева обсужден на заседании семинара отдела математического программирования 9 сентября 2016 г., протокол № 724.

Заведующий отделом математического программирования,
д. ф.-м. н., проф. РАН

М.Ю.Хачай



Handwritten signature of O.N. Ulyanov
Ульянов О.Н.

12 сентября 2016

Федеральное государственное бюджетное учреждение науки
Институт математики и механики им. Н.Н. Красовского
Уральского отделения Российской академии наук (ИММ УрО РАН)

620990, Российская Федерация, г. Екатеринбург, ул. Софьи Ковалевской, д. 16;

телефон: +7 (343) 374-83-32

тел./факс: +7 (343) 374-25-81

e-mail: dir-info@imm.uran.ru

Web: <http://www.imm.uran.ru>