

ОТЗЫВ

официального оппонента о диссертации Виноградова Дмитрия Вячеславовича «Вероятностно-комбинаторный формальный метод обучения, основанный на теории решеток», представленной на соискание ученой степени доктора физико-математических наук по специальности 05.13.17 «Теоретические основы информатики»

Многие задачи интеллектуального управления требуют извлечения знаний из обучающих выборок. В некоторых случаях достаточно соотнесение ситуации к одному из нескольких классов. В этих случаях достаточно воспользоваться хорошо развитой в настоящее время техникой нейронных сетей. Однако гораздо чаще лицу, принимающему решение, требуется обоснование принятого решения, которое невозможно получить в нейросетевой парадигме. Такого рода объяснения принимаемых решений порождают алгоритмы комбинаторного анализа данных. Они, в свою очередь, обладают высокой вычислительной сложностью и не могут справиться с обучающими выборками большого объема. Наиболее известным методом комбинаторного анализа данных является ДСМ-метод порождения гипотез, созданный и развиваемый отечественной школой исследователей под руководством проф. В.К. Финна. Попытка расширить технологию ДСМ-метода на большие выборки представляется очень актуальной, особенно из-за взрывного объема обрабатываемых данных. Этой актуальной задаче и посвящено диссертационное исследование Д.В. Виноградова.

Диссертационная работа Д.В. Виноградова посвящена созданию вероятностно-комбинаторного (ВКФ-)метода машинного обучения и математическому исследованию алгоритмов и их ключевых параметров.

Диссертация Д.В. Виноградова соответствует традиционной для математических диссертаций схеме:

- формулируются или напоминаются ключевые определения;
- доказываются вспомогательные леммы;
- доказываются ключевые теоремы;
- дается неформальное пояснение смысла доказанных результатов.

Первая глава определяет необходимые понятия и базовые алгоритмы из прикладной теории решеток. Сначала автор вводит операции «Замыкай-по-одному». Потом, после анализа сложности этих операций предлагается ленивая схема их вычислений, и оценивается эффективность замены классического варианта ленивым. Наконец, излагается алгоритм кодирования объектов битовыми строками (с заменой операции сходства на побитовое умножение), и доказываются его корректность.

Вторая глава стоит несколько в стороне от основной линии исследований: она посвящена изложению результатов автора о проблеме переобучения. Доказываются теоремы о невозможности устранить переобучение ни через границу на минимальное число родителей гипотез, ни «запретом контр-примеров», т.е. примеров, не обладающих целевым свойством. Во второй половине главы 2 получается явный вид производящие

функции вероятностей преодолеть нужное число контр-примеров сходством заданного числа родителей.

Третья глава самая объемная: она вероятностным алгоритмам для поиска сходства. Все они соответствуют цепям Маркова на решетке сходств. Исследованы немонотонная, монотонная, спаривающая, ленивая спаривающая, остановленная ленивая спаривающая цепи Маркова. Для спаривающих цепей Маркова доказаны все ключевые свойства: останавливаемость с вероятностью единица, экспоненциальное убывание расстояния между распределениями остановки между остановленным и неостановленным вариантами в метрике тотальной вариации. В случае Булеана удалось получить оценку на среднее время нахождения случайного подмножества спаривающей цепью Маркова, отличающейся от прямого задания его последовательностью испытаний Бернулли на логарифмический множитель. Более того, доказана и теорема о сильной концентрации времени работы около своего среднего. Для монотонной цепи Маркова в случае Булеана доказаны теоремы об оценке времени перемешивания и точности этой оценки.

Четвертая глава посвящена изложению ключевых процедур: обобщения обучающих примеров в гипотезах и предсказания целевого свойства у новых примеров по аналогии. Для главного параметра – числа необходимых гипотез – доказана теорема об оценке числа гипотез, необходимых для надежного предсказания всех важных положительных примеров. Во второй части главы 4 описывается программная реализация предлагаемого подхода и даны результаты применения этой программной системы к двум массивам реальных данных. Предложенный автором подход продемонстрировал преимущества над известным алгоритмом комбинаторного обучения (покрытие с помощью целочисленного обучения) и возможность его применения к массивам данных, содержащим несколько тысяч обучающих примеров.

Новизна, теоретическая и практическая значимость результатов выполненного Д.В. Виноградовым исследования позволяет квалифицировать его диссертационную работу как открытие нового научного направления в теоретической информатике – «машинное обучение, с использованием методов теории решеток».

Оформление диссертационной работы соответствует требованиям ВАК РФ. Автореферат полностью отражает содержание диссертации. Все положения диссертационной работы достоверны, а выводы и заключения математически обоснованы. Доказанные в работе математические результаты приводят или к алгоритмическим улучшениям, или к оценке ключевых параметров. Работоспособность предложенного автором подхода подтверждена применением ее программной реализации к реальным массивам из репозитория данных для тестирования алгоритмов машинного обучения.

Результаты диссертационного исследования опубликованы в 23 статьях, в том числе – в 17 работах в изданиях, рекомендованных ВАК РФ для публикации основных научных результатов диссертаций на соискание ученой степени доктора наук. Полученные автором результаты были представлены им на отечественных и международных конференциях и научных семинарах.

Материал диссертации хорошо структурирован, изложение характеризуется логической последовательностью. Текст написан, в основном, хорошим языком, стиль изложения четкий и понятный.

О недостатках работы.

1. Теорема 1.1 нуждается в уточнении формулировки. Текущее содержание формулировки хотя и является абсолютно корректным, тем не менее требует от читателя нетривиального умозаключения об алгоритмическом значении этого результата.
2. Из текста не понятно, как осуществлялось кодирование обучающих и тестовых примеров битовыми строками при описании проведенных экспериментов в главе 4.

Перечисленные недостатки не умаляют значимости диссертационной работы Виноградова Д.В. и не оказывают влияния на ее общую положительную оценку.

Диссертационная работа Д.В. Виноградова «Вероятностно-комбинаторный формальный метод обучения, основанный на теории решеток» полностью удовлетворяет критериям, установленным «Положением о присуждении ученых степеней», утвержденного Постановлением Правительства РФ от 24 сентября 2013 г. № 842.

Данная работа соответствует всем требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени доктора физико-математических наук, а ее автор – Виноградов Дмитрий Вячеславович – заслуживает присуждения ему ученой степени доктора физико-математических наук по специальности 05.13.17 – «Теоретические основы информатики».

Официальный оппонент:

Академик РАН, д.ф.-м.н., профессор,
главный научный сотрудник

«23» января 2019 года

Почтовый адрес:

117997, Москва, ул. Профсоюзная, д. 65, ИПУ РАН

Телефон: +7(495)330-55-01

Организация – место работы: Федеральное государственное бюджетное учреждение науки «Институт проблем управления им. В.А. Трапезникова Российской академии наук»

Должность: главный научный сотрудник

Адрес электронной почты: yassilyev_sn@mail.ru

Веб-сайт: <http://www.ipu.ru>

С.Н. Васильев



Подпись *Васильев С.Н.*
ЗАВЕРЯЮ
ЗАВ. ОТДЕЛОМ КАДРОВ
И.А. ГАВРИЛОВА
Тач-