

«Утверждаю»

Зам. директора Института проблем
управления им. В.А. Трапезникова РАН



И.Н.Барабанов

«29» сентября 2014 г.

ОТЗЫВ

ведущей организации — Института проблем управления
им. В.А.Трапезникова РАН
на диссертационную работу Толстихина Ильи Олеговича
на тему **“Неравенства концентрации вероятностной меры
в трансдуктивном обучении и РАС-байесовском анализе”**,
представленную на соискание ученой степени кандидата
физико-математических наук по специальности
05.13.17 – «Теоретические основы информатики».

Актуальность темы

Задачи поиска закономерностей или восстановления функциональных зависимостей в наблюдаемых данных сегодня играют ключевую роль во многих прикладных областях. При этом важным направлением исследований является теоретический анализ свойств процедуры обучения, который нередко ведет к новым более точным алгоритмам обучения. Строгий вероятностный подход к изучению качества работы процедур обучения был сформулирован в конце 1960-х годов в работах В.Н.Вапника и А.Я.Червоненкиса, которые позже стали классическими и легли в основу современной теории статистического обучения. Именно ей посвящена диссертационная работа И.О.Толстихина. Основная задача, изучаемая статистической теорией обучения, заключается в минимизации значения функционала среднего риска по выбранному классу измеримых отображений. Поскольку совместное распределение данных заранее неизвестно, эту задачу решают на основе случайной обучающей выборки, минимизируя эмпирический риск на ней. Минимальный размер обучающей выборки, необходимый для достижения желаемой точности решения, определяется при этом с помощью вероятностных оценок для минимума эмпирического риска, известных как оценки обобщающей способности.

Несмотря на более чем 50-летнюю историю попыток улучшения точности оценок в теории статистического обучения, они по-прежнему остаются чрезмерно завышенными для применения на практике. Поэтому тема диссертационной работы И.О.Толстихина, посвященной их дальнейшему уточнению на основе современных результатов неравенств концентрации вероятностной меры, без сомнений является актуальной и представляет большой интерес не только с теоретической, но и с практической точки зрения. Стоит также отметить, что задачу минимизации среднего риска можно рассматривать как частный случай задачи стохастической оптимизации. Эта важная связь между двумя активно развивающимися

направлениями математики, широко обсуждаемая в работах последних лет, придает дополнительный интерес к рассматриваемым в диссертационной работе задачам.

Содержание

Диссертационная работа состоит из введения, шести глав и заключения. В первой главе приводится подробный обзор современного состояния теории неравенств концентрации меры для независимых случайных величин. Вторая глава посвящена неравенствам концентрации для выборок без возвращения, и в ней получено два новых неравенства концентрации для эмпирических процессов. В третьей главе представлен подробный обзор классических и современных результатов теории статистического обучения (в стандартной постановке обучения с учителем). Четвертая глава рассматривает специальную постановку так называемого трансдуктивного обучения, и для нее на основе полученных во второй главе результатов доказаны новые оценки обобщающей способности и избыточного риска. В пятой главе рассматривается комбинаторная теория переобучения, и на основе оригинальной идеи учета орбит разбиений выборок получены новые оценки обобщающей способности для трех модельных семейств отображений. Наконец, в шестой главе рассматривается РАС-Байесовский подход к оцениванию среднего риска, и получено новое РАС-Байесовское неравенство типа Бернштейна.

Основные результаты и их новизна

В рамках диссертационной работы И.О.Толстихина получены следующие новые результаты:

(1) Рассмотрены эмпирические процессы для выборок без возвращения, и для них получено два новых неравенства концентрации, учитывающие дисперсии случайных величин. Благодаря учету дисперсий новые неравенства в ряде интересных случаев позволяют получать гораздо более точные результаты по сравнению с известными ранее неравенствами типа МакДиармида.

(2) Для задачи трансдуктивного обучения получены новые оценки избыточного риска и обобщающей способности на основе так называемого локального подхода. Новая оценка избыточного риска учитывает локальную геометрическую структуру семейства функций, что в ряде случаев ведет к ее хорошей точности.

(3) В рамках комбинаторной теории переобучения работа И.О.Толстихина предлагает новую идею учета орбит разбиений генеральной выборки при вычислении вероятности переобучения, которая во многих случаях позволяет упростить вычисление вероятности переобучения. С помощью этой идеи получены новые оценки вероятности переобучения для трех различных модельных семейств функций.

(4) В рамках РАС-Байесовского анализа получено новое неравенство для среднего риска рандомизированного алгоритма прогнозирования, учитывающее дисперсии потерь. В отличие от известных ранее аналогов, новое неравенство не требует знания неизвестных параметров распределения и зависит только от наблюдаемых величин.

Все результаты, выносимые на защиту в диссертации, впервые получены диссертантом.

Достоверность полученных результатов

Математические результаты диссертационной работы И.О.Толстихина сформулированы в виде теорем с исчерпывающими и корректными доказательствами. Достоверность и обоснованность теоретических результатов обеспечивается подробным изложением доказательств, допускающих их проверку. Достоверность и обоснованность результатов экспериментов и выводов, сформулированных на основе них в диссертации, определяется подробными описаниями всех экспериментов, обеспечивающими их воспроизводимость.

Значимость полученных результатов

Диссертационная работа И.О.Толстихина в первую очередь имеет теоретический характер. В работе получено множество результатов, включая локальную оценку избыточного риска для задач трансдуктивного обучения и новое РАС-Байесовское неравенство Бернштейновского типа, которые, безусловно, способствуют дальнейшему развитию теории статистического обучения. Кроме того стоит отметить, что новые неравенства концентрации типа Талаграна, полученные для эмпирических процессов и выборок без возвращений, являются достаточно общими математическими результатами и могут быть интересны не только в теории статистического обучения, но и за ее пределами. Практическая значимость диссертационной работы заключается в том, что минимизация полученного в ней РАС-Байесовского неравенства является универсальным способом получения новых алгоритмов машинного обучения, имеющих гарантированно высокую обобщающую способность. Результаты работы И.О.Толстихина могут найти применение в научных учреждениях и компаниях, деятельность которых связана с анализом данных и приложениями машинного обучения.

Замечания

Существенных замечаний по диссертации нет.

(1) Название диссертации звучит несколько «заумно» и рассчитано на узких специалистов. Термины «трансдуктивное обучение» и «РАС-Байесовский анализ» появляются в середине или конце диссертации, а в автореферате вообще не разъяснены. Гораздо естественнее было бы назвать работу «Уточнение оценок в статистической теории обучения».

(2) Автор хорошо знаком с публикациями по теме диссертации; список литературы достаточно полон. Однако естественно было бы ссылки на российских авторов давать на русском языке. Это относится прежде всего к таким основополагающим работам, как труды Вапника-Червоненкиса и Бернштейна (все они почему-то даны по-английски).

(3) С.Н.Бернштейну «не повезло» и в другом месте. На стр.12 говорится «Первые исследования были связаны с суммами случайных величин, для которых начиная с 1950-х годов С. Н. Бернштейном, Дж. Беннеттом, В. Хефдингом и другими авторами были получены верхние оценки...» На самом деле неравенство Бернштейна относится к 1911 г.

Однако, приведенные выше замечания являются редакционными и не могут повлиять на общую положительную оценку диссертационной работы.

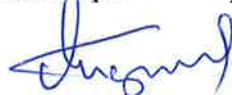
Заключительная оценка

Диссертационная работа Толстихина Ильи Олеговича является законченной научно-квалификационной работой, имеющей теоретическую и практическую ценность. Представленные в работе результаты и выводы корректно обоснованы. Результаты диссертационной работы своевременно и полно опубликованы в 7 печатных трудах, 3 из которых — в изданиях, рекомендованных ВАК РФ. Результаты докладывались автором на ведущих международных конференциях. Автореферат правильно и полно отражает содержание диссертационной работы.

Диссертация удовлетворяет всем требованиям «Положения о порядке присуждения ученых степеней», утвержденного Постановлением Правительства РФ от 24 сентября 2013 года № 842, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 – «Теоретические основы информатики», а ее автор — Толстихин Илья Олегович — заслуживает присуждения искомой ученой степени.

Настоящий отзыв обсужден и одобрен на заседании научного семинара Лаборатории 7 «Адаптивных и робастных систем» им. Я.З.Цыпкина Федерального государственного бюджетного учреждения науки Института проблем управления им. В.А.Трапезникова Российской академии наук (протокол № 1 от 23 сентября 2014 г.)

Зав. лабораторией №7 ИПУ РАН д.ф.-м.н.



/М.В.Хлебников/

Гл. научн. сотр. лаборатории №7 ИПУ РАН д.т.н.



/Б.Т.Поляк/

23 сентября 2014 года