

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА
на диссертационную работу Бахтеева Олега Юрьевича
**«Байесовский выбор субоптимальной структуры модели глубокого
обучения»**,
представленную на соискание ученой степени
кандидата физико-математических наук по специальности
05.13.17 – «Теоретические основы информатики».

Актуальность проведенного в диссертации исследования. Диссертационная работа О.Ю. Бахтеева посвящена исследованию актуальной фундаментальной проблемы выбора структуры модели глубокого обучения. Основная трудность задачи, поставленной и решенной в диссертации, состоит в том, что оптимизируемая модель обладает большим числом параметров и высокой вычислительной сложностью. Подход к решению задачи, выбранный автором, основан на статистических критериях сложности модели, исследованию этих проблем посвящены многочисленные работы отечественных и зарубежных ученых (В.В. Стрижов, К.В. Воронцов, К. Бишоп, Д. МакКай и др.) В оппонируемой работе этот подход получил дальнейшее развитие; в ней предлагается заменить исходную оптимизационную задачу на приближенную, решение которой позволяет снизить вычислительную сложность за счет допустимой потери в точности, а также разрабатывается алгоритм построения модели субоптимальной сложности и оптимизации параметров модели. **Актуальность диссертационной работы О.Ю.Бахтеева не вызывает сомнений**, поскольку в результате проведенных в ней исследований преодолевается существенная трудность решения задачи глубокого обучения, состоящая в решении невыпуклых оптимизационных задач с большим числом параметров.

Научная новизна исследований. В диссертации предложен новый метод выбора моделей глубокого обучения оптимальной и субоптимальной сложности. Сложность модели определяется средним значением правдоподобия выборки, усредненном по случайным параметрам модели и ее структуры, называемом обоснованностью модели. Для оценки сложности даются методы вычисления новых нижних оценок обоснованности модели, изучаются их свойства.

Предложен новый метод оптимизации модели глубокого обучения выбором ее параметров и гиперпараметров (вводятся вероятностные

предположения о распределениях параметров и структуры модели, под гиперпараметрами понимаются параметры этих распределений): на основе байесовского вариационного вывода проводится градиентная оптимизация параметров и гиперпараметров; критерием оптимизации служит обобщенная функция обоснованности модели. Проанализированы свойства задачи оптимизации структуры и параметров модели при различных значениях метапараметров, а также связь предлагаемой оптимизационной задачи с другими задачами выбора моделей

Достоверность и обоснованность исследований представленной работы подтверждается строгими доказательствами математических утверждений, полученных в рамках диссертационного исследования.

Практическая значимость работы состоит в разработке реализуемых алгоритмов оптимизации параметров модели глубокого обучения и применении полученных моделей для решения прикладных задач глубокого обучения – задач регрессии, классификации и др.

Диссертационная работа состоит из введения, пяти глав и заключения.

В первой главе даются основные определения, вводятся понятия и обозначения, используемые в диссертации. Описана графовая математическая модель глубокого обучения, дана ее вероятностная интерпретация, введена ее многоуровневая параметризация. Описаны решаемые в работе задачи, в частности, дано формальное описание задачи выбора модели глубокого обучения как двухуровневой задачи оптимизации.

Во второй главе предложены критерии оптимальности и субоптимальности моделей. Рассмотрена задача выбора моделей глубокого обучения. Предложен метод получения вариационной нижней оценки обоснованности модели на основе градиентного спуска и его модификаций.

В третьей главе рассмотрена задача оптимизации гиперпараметров моделей. Исследованы методы оптимизации на основе градиентных методов на выборках большой мощности. Показано, что при большом количестве гиперпараметров, сравнимым с количеством параметров модели, рассматриваемые алгоритмы предпочтительнее стохастических.

В четвертой главе рассмотрена задача выбора структуры модели глубокого обучения. Определена параметрическая сложность модели. Доказана теорема о связи параметрической сложности модели и относительной вариационной плотности параметров модели. Определена обобщающая задача. Доказана теорема о виде обобщающей задачи. Доказаны теоремы об оценках параметрической сложности при различных значениях метапараметров.

В пятой главе описаны результаты экспериментальных исследований и сравнительного анализа предложенных методов в применении к задачам классификации и регрессии с использованием моделей глубокого обучения. Таким образом, диссертационная работа О.Ю. Бахтеева является законченным научным исследованием, в котором получены следующие **основные результаты**.

1. Сформулировано формальное графовое описание структуры модели глубокого обучения.
2. Разработан алгоритм выбора параметров, гиперпараметров и структуры модели глубокого обучения.
3. Предложены критерии оптимальности и субоптимальности модели глубокого обучения.
4. Поставлена и решена задача оптимизации модели глубокого обучения, соответствующая байесовскому выбору модели.

Среди **замечаний** к работе можно выделить следующие.

1. В работе имеется значительное число опечаток, в основном связанное с несогласованностью окончаний слов (в тексте на стр. 12, 15, 16, 17, 27, 32, 49, в подписях к рисункам 2.1 на стр. 38, к рис. 3.5 на стр.57 и др.)
2. Обзор публикаций по теме диссертации, по сути, сводится к перечислению работ тех или иных авторов, было бы полезно указать недостатки и достоинства методов и обосновать тот подход, который выбран диссертантом.
3. Теоремы, представленные в главе 4, имеют общие условия, которые можно было бы сформулировать отдельно для ясности изложения.
4. Отсутствует анализ сходимости предложенного метода выбора структуры модели.
5. Отсутствуют рекомендации по назначению метапараметров в задаче оптимизации.

Эти замечания носят редакционный характер, не являются принципиальными и не влияют на главные теоретические и практические результаты диссертационной работы О.Ю. Бахтеева.

Результаты диссертационной работы **изложены в одиннадцати статьях, из которых девять входят в перечень ВАК РФ**. Результаты докладывались на российских и международных конференциях по анализу данных и исследованию операций.

Автореферат адекватно отражает содержание диссертационной работы.

Диссертация соответствует критериям, установленным Положением о присуждении ученых степеней. Диссертация обладает внутренним единством, содержит новые научные результаты и рекомендации по их практическому использованию. В ней содержится решение актуальной проблемы оптимизации модели глубокого обучения.

Работа полностью соответствует требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 «Теоретические основы информатики», а ее автор, **О.Ю. Бахтеев, заслуживает присуждения ему ученой степени кандидата наук по данной специальности.**

Официальный оппонент:

д.ф.-м.н., профессор,
заведующий кафедрой математического моделирования и информатики
физического факультета
ФГБОУ ВО «Московского государственного
университета имени М.В.Ломоносова»

17.01.2020



Чуличков Алексей Иванович

Контактные данные:

тел.: +7 (903) 570-75-50, e-mail: achulichkov@physics.msu.ru

Специальность, по которой официальным оппонентом защищена диссертация:
05.13.18 - Математическое моделирование, численные методы и комплексы программ (физ.-мат. науки)

Адрес места работы:

119991, г. Москва, Ленинские горы, д.1, стр.2, ФГБОУ ВО «Московский государственный университет имени М.В.Ломоносова», физический факультет, кафедра математического моделирования и информатики
Тел.: +7 (495) 939-13-32; e-mail: achulichkov@physics.msu.ru

Подпись зав. кафедрой ФГБОУ ВО «Московского государственного университета имени М.В.Ломоносова» А.И. Чуличкова удостоверяю:

Декан физического факультета МГУ имени М.В. Ломоносова
профессор
доктор физико-математических наук, профессор



Н.Н. Сысоев