

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»

Оценка актуальности темы диссертационной работы

Проблематика диссертационной работы. Работа посвящена исследованию задачи выбора модели глубинного обучения и оптимизации параметров и структуры этой модели. Эта задача является актуальной, так как автоматический и адекватный выбор модели позволяет эффективно улучшать ее свойства, такие как качество и время работы. В работе предложено формальное описание структуры модели глубинного обучения, а также подход к выбору модели, основанный на вероятностных предположениях. Использование вероятностного представления разработанных подходов позволяет естественным образом описывать различные варианты постановок задач и приводит к задаче оптимизации, обобщающей ряд критериев выбора модели. Проведено исследование свойств этой задачи.

Актуальность. Диссертационная работа является актуальной: разработанные методы могут быть использованы для решения задач выбора модели в задачах машинного обучения, сходные постановки задач активно исследуются в последние годы.

Структура и содержание работы

Диссертационная работа состоит из введения, пяти глав, заключения и списка литературы.

Во введении обоснована актуальность работы, определены цели и задачи исследования.

В первой главе введены основные понятия и обозначения. Предложено графовое описание структуры модели, введен необходимый вероятностный аппарат, сформулирована задача выбора модели.

Во второй главе предложен метод выбора модели на основе вариационного вывода. Предложен метод получения вариационной нижней оценки на основе градиентных методов оптимизации моделей. Доказана теорема об энтропии распределения, полученного под действием оператора градиентного спуска. Путем проведения численных экспериментов сделан анализ методов выбора модели.

В третьей главе рассмотрена двухуровневая задача оптимизации гиперпараметров модели. Исследованы различные градиентные методы оптимизации гиперпараметров модели.

В четвертой главе рассмотрена задача выбора структуры модели. Введены вероятностные предположения о структуре модели глубинного обучения.

Предложена обобщающая задача выбора модели, проведено ее сопоставление с другими задачами выбора модели. Исследованы асимптотические свойства предложенной задачи оптимизации.

В пятой главе приведены результаты экспериментального анализа свойств предложенных алгоритмов и сравнения предложенных в работе алгоритмов с существующими.

Основные результаты и их научная новизна

1. Предложены критерии оптимальной и субоптимальной сложности модели глубинного обучения. Предложен метод выбора модели, основывающийся на данных критериях.
2. Предложен метод графового описания структуры модели глубинного обучения. Предложена задача оптимизации структуры модели, обобщающая ряд методов выбора модели. Исследованы ее свойства.
3. Предложен и исследован метод получения вариационной нижней оценки обоснованности модели глубинного обучения на основе градиентных методов оптимизации.
4. Рассмотрены практические задачи выбора модели глубинного обучения, а также оптимизации ее параметров и гиперпараметров.

Новизна. Предложены методы выбора структуры модели глубинного обучения на основе Байесовского вывода. Метод выбора структуры модели, основанный на оптимизации вариационной нижней оценки, является новым. Также является новым предложенный обобщенный метод выбора структуры модели.

Достоверность полученных результатов. Достоверность подтверждается публикациями результатов в высокоцитируемых научных журналах, а также докладами на российских и международных конференциях по анализу данных и исследованию операций. Научные положения, выводы и рекомендации, сформулированные в диссертации, являются обоснованными и подкреплены необходимыми математическими утверждениями.

Замечания

1. Представление результатов можно было бы улучшить. Например, обзор работ, рассматривающих понятия сложности модели глубинного обучения дублируется на странице 4 и странице 17.

2. В главе 2 отсутствует анализ требуемого количества оптимизаций моделей для получения адекватной вариационной оценки обоснованности модели.
3. Анализ свойств предложенной задачи оптимизации в главе 4 был проведен в достаточно строгих предположениях об области определения задачи. Можно ли сделать данные предположения более реалистичными?
4. В главе 5 описание разработанного комплекса программ не является достаточно полным
5. Использование Theano и Matlab в качестве библиотеки и среды для построения моделей глубинного обучения соответственно затрудняет переиспользование полученных результатов, так как оба инструмента не являются наиболее популярными в современном академическом и индустриальном сообществах исследователей моделей машинного и глубинного обучения. В целом релевантность раздела 5.1 решаемой в диссертационной работе задачи не обоснована.

Заключительная оценка

Приведенные замечания не снижают общей высокой оценки диссертационной работы.

Результаты диссертационной работы изложены в 11 научных статьях, из которых 9 опубликованы в журналах из Перечня ведущих рецензируемых научных журналов, включенных Высшей аттестационной комиссией России в список изданий, рекомендуемых для опубликования основных научных результатов диссертации на соискание ученой степени кандидата и доктора наук. Результаты докладывались на Всероссийских и международных конференциях. Автореферат достаточно полно и правильно отражает содержание диссертации.

Диссертационная работа представляет собой законченное научное исследование, содержит решение задачи, имеющей значение для развития методов анализа данных, она написана автором самостоятельно, обладает внутренним единством, содержит новые научные результаты, полученные автором лично, и свидетельствует о личном вкладе автора в науку. В диссертации приведены рекомендации по использованию научных выводов.

Работа полностью соответствует требованиям ВАК РФ (п. 9 «Положения о присуждении 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, О.Ю. Бахтеев, заслуживает

присуждения ему ученой степени кандидата физико-математических наук по данной специальности.

Официальный оппонент

Кандидат физико-математических наук,
руководитель лаборатории,

Центр по научным и инженерным вычислительным
технологиям для задач с большими массивами данных
Сколковского института науки и технологий

<http://www.skoltech.ru/>

Адрес: 143026 Москва, Территория Инновационного Центра “Сколково”, улица Нобеля,
д. 3

Тел.: +7 (495) 280 14 81

E-mail: a.zaytsev@skoltech.ru

Зайцев Алексей Алексеевич

16 января 2020 г.

Юлия Зайцева А.А. подтверждает.
РУКОВОДИТЕЛЬ ОТДЕЛА
КАДРОВОГО АДМИНИСТРИРОВАНИЯ

