

УТВЕРЖДАЮ



Проректор по научной работе и
программам развития, к.ф.- м.н.

Баган Виталий
Анатольевич

«10» октября 2019 г.

ЗАКЛЮЧЕНИЕ

федерального государственного автономного образовательного
учреждения высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)» (МФТИ)

Диссертация «Байесовский выбор субоптимальной структуры модели глубокого обучения» выполнена на кафедре интеллектуальных систем федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)» Министерства науки и высшего образования Российской Федерации.

В период подготовки диссертации с 2016 г. по 2019 г. соискатель **Бахтеев Олег Юрьевич** работал в отделе исследований акционерного общества «Антиплагиат» в должности разработчика-исследователя.

В 2016 г. соискатель **Бахтеев Олег Юрьевич** с отличием окончил федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (государственный университет)» по направлению подготовки 03.04.01 Прикладные математика и физика. В 2016 г. поступил в очную аспирантуру МФТИ. Справка об обучении в аспирантуре и сдаче кандидатских экзаменов выдана в 2019 г. в федеральном государственном автономном образовательном учреждении высшего образования «Московский физико-технический институт (национальный исследовательский университет)».

Научный руководитель – доктор физико-математических наук Стрижов Вадим Викторович. Основное место работы – Вычислительный центр им. А.А. Дородницына Российской академии наук Федерального исследовательского центра «Информатика и управление» Российской академии наук, отдел интеллектуальных систем, ведущий научный сотрудник.

По итогам обсуждения диссертации «Байесовский выбор субоптимальной структуры модели глубокого обучения» *принято следующее заключение.*

Тема диссертации является актуальной для развития методов выбора моделей глубокого обучения. Предложены критерии субоптимальности и оптимальности модели, а также метод выбора модели субоптимальной сложности.

Цели и задачи диссертации:

1. Исследовать методы построения моделей глубокого обучения оптимальной и субоптимальной сложности.
2. Предложить критерии оптимальной и субоптимальной сложности модели глубокого обучения.
3. Предложить метод выбора субоптимальной структуры модели глубокого обучения.
4. Предложить алгоритм построения модели субоптимальной сложности и оптимизации ее параметров.

Основные результаты диссертации:

1. Предложен метод байесовского выбора оптимальной и субоптимальной структуры модели глубокого обучения с использованием автоматического определения релевантности параметров.
2. Предложены критерии оптимальной и субоптимальной сложности модели глубокого обучения.
3. Предложен метод графового описания моделей глубокого обучения.
4. Предложено обобщение задачи оптимизации структуры модели, включающее ранее описанные методы выбора модели: оптимизация обоснованности модели, последовательное увеличение сложности модели, последовательное снижение сложности модели, полный перебор вариантов структуры модели.
5. Предложен метод оптимизации вариационной оценки обоснованности модели на основе метода мултистарта задачи оптимизации.
6. Предложен алгоритм оптимизации параметров, гиперпараметров и структурных параметров моделей глубокого обучения.
7. Исследованы свойства оптимизационной задачи при различных значениях метапараметров. Рассмотрены ее асимптотические свойства.

Все *результаты диссертации получены лично соискателем* при научном руководстве д.ф.-м.н. Стрижова В.В.

Научная новизна работы заключается в том, что разработан новый подход к построению моделей глубокого обучения. Предложены критерии субоптимальной и оптимальной сложности модели, а также исследована их связь. Предложен метод построения модели глубокого обучения субоптимальной сложности. Исследованы методы оптимизации гиперпараметров и параметров модели. Предложена обобщенная задача выбора модели глубокого обучения.

Практическая ценность результатов диссертации заключается в том, что предложенные в работе методы предназначены для построения моделей глубокого обучения в задачах регрессии и классификации; оптимизации гиперпараметров полученной модели; выборе модели из конечного множества заданных моделей; получения оценок переобучения модели.

Обоснованность и достоверность результатов и выводов

Достоверность результатов подтверждена математическими доказательствами, экспериментальной проверкой полученных методов на реальных задачах; публикациями результатов исследования в рецензируемых научных изданиях, в том числе рекомендованных ВАК. Результаты работы докладывались и обсуждались на российских и международных научных конференциях по анализу данных и исследованию операций.

Материалы диссертации опубликованы автором достаточно полно в следующих работах:

1. Грабовой А.В., Бахтеев О.Ю., Стрижов В.В. Определение релевантности параметров нейросети. // Информатика и её применения. – 2019. –13:2. – С.62-71.
2. Смердов А.Н., Бахтеев О.Ю., Стрижов В.В. Выбор оптимальной модели рекуррентной сети в задачах поиска парафраза. // Информатика и ее применения. – 2018. – 12:4 – С.63–69.
3. Огальцов А.В., Бахтеев О.Ю. Автоматическое извлечение метаданных из научных PDF-документов. // Информатика и её применения. – 12:2. – 2018. – С.75–82.
4. Бахтеев О.Ю., Стрижов В.В. Выбор моделей глубокого обучения субоптимальной сложности. // Автоматика и телемеханика. – 2018. – No8. – С. 129-147.
5. Bakhteev, O., Khazov, A. Author masking using sequence-to-sequence models: Notebook for PAN at CLEF 2017 //CEUR Workshop Proceedings. – 2017.
6. Romanov, A.V., Kuznetsova, M.V., Bakhteev, O.Yu., Khritankov, A.S. Machine-translated text detection in a collection of Russian scientific papers // Computational Linguistics and Intellectual Technologies. – 2016.
7. Бахтеев О.Ю., Попова М.С., Стрижов В.В. Системы и средства глубокого обучения в задачах классификации. // Системы и средства информатики. – 2016. – No 26.2. – С. 4-22.
8. Bakhteev, O., Kuznetsova, R., Romanov, A., Khritankov, A. A monolingual approach to detection of text reuse in Russian-English collection. // Proceedings of Artificial Intelligence and Natural Language and Information Extraction, Social Media and Web Search FRUCT Conference, AINL-ISMW FRUCT. – 2015

Личный вклад соискателя в работах с соавторами заключается в следующем:

[4,5,7,8] – постановка задачи и подготовка теоретической части, проведение вычислительных экспериментов; [1,2,3] – постановка задачи и подготовка теоретической части, [6] – анализ и обработка результатов экспериментов.

Основные результаты работы докладывались на следующих научных конференциях и семинарах:

1. “Variational learning across domains with triplet information”, Международная конференция «Visually Grounded Interaction and Language workshop, Conference on Neural Information Processing Systems», 2018.
2. “Байесовский выбор наиболее правдоподобной структуры модели глубокого обучения”, Международная конференция «Интеллектуализация обработки информации», 2018.
3. “ParaPlagDet: The system of paraphrased plagiarism detection”, Международная конференция «Big Scholar at conference on knowledge discovery and data mining», 2018.

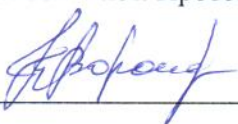
4. “Детектирование переводных заимствований в текстах научных статей из журналов, входящих в РИНЦ”, Всероссийская конференция «Математические методы распознавания образов ММРО», 2017.
5. “Градиентные методы оптимизации гиперпараметров моделей глубокого обучения”, Всероссийская конференция «Математические методы распознавания образов ММРО», 2017.
6. “Author Masking using Sequence-to-Sequence Models”, Международная конференция «Conference and Labs of the Evaluation Forum», 2017.
7. “Machine-Translated Text Detection in a Collection of Russian Scientific Papers”, Международная конференция по компьютерной лингвистике и интеллектуальным технологиям «Диалог-21», 2017.
8. “Выбор модели глубокого обучения субоптимальной сложности с использованием вариационной оценки правдоподобия”, Международная конференция «Интеллектуализация обработки информации», 2016.
9. “A monolingual approach to detection of text reuse in Russian-English collection”, Международная конференция «Artificial Intelligence and Natural Language Conference», 2015.
10. “Восстановление панельной матрицы и ранжирующей модели в разнородных шкалах”, Всероссийская конференция «57-я научная конференция МФТИ», 2014.

Содержание диссертации соответствует паспорту специальности 05.13.17 – теоретические основы информатики, в частности, пунктам:

- п. 5 – Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечения; разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений.
- п. 6 – Разработка методов, языков и моделей человекомашинного общения; разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения данных из текстов на естественном языке.

Диссертация «Байесовский выбор субоптимальной структуры модели глубокого обучения» **Бахтеева Олега Юрьевича** рекомендуется к защите на соискание ученой степени кандидата физико-математических наук по специальности **05.13.17 – теоретические основы информатики**.

Заключение принято на заседании кафедры интеллектуальных систем МФТИ. Присутствовало на заседании 13 человек. Результаты голосования: «за» – 13 чел., «против» – нет, «воздержались» – нет. Протокол № 1 от 9 октября 2019 года.



Воронцов Константин Вячеславович
профессор РАН, д.ф.-м.н.,
заместитель заведующего кафедрой
интеллектуальных систем МФТИ