

**ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА
Д 212.268.05, СОЗДАННОГО НА БАЗЕ ФЕДЕРАЛЬНОГО
ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО
УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ «ТОМСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ
И РАДИОЭЛЕКТРОНИКИ» (ТУСУР) МИНИСТЕРСТВА НАУКИ
И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ
КАНДИДАТА НАУК**

аттестационное дело № _____

решение диссертационного совета от 24 декабря 2020 г. № 4

О присуждении Грибкову Егору Игоревичу, гражданину Российской Федерации, учёной степени кандидата технических наук.

Диссертация «Нейросетевые модели на основе системы переходов для извлечения структурированной информации о продуктах из текстов пользователей» по специальности 05.13.17 – «Теоретические основы информатики» принята к защите 23 октября 2020 г. (протокол № 3) диссертационным советом Д 212.268.05 на базе ТУСУРа (634050, г. Томск, пр. Ленина, 40). Приказ о создании диссертационного совета № 1236/нк от 12.10.2015 г.

Соискатель Грибков Егор Игоревич, 1993 года рождения, в 2017 г. окончил ТУСУР. С 2016 по 2020 г. обучался в аспирантуре ТУСУРа. Работает младшим научным сотрудником в лаборатории объектно-ориентированного моделирования информационных систем кафедры автоматизации обработки информации (АОИ) ТУСУРа.

Диссертация выполнена на кафедре АОИ ТУСУРа.

Научный руководитель — доктор технических наук профессор Ехлаков Юрий Поликарпович, профессор кафедры АОИ ТУСУРа.

Официальные оппоненты: Спицын Владимир Григорьевич, доктор технических наук, профессор, профессор отделения информационных технологий инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета; Пимонов Александр Григорьевич,

доктор технических наук, профессор, заведующий кафедрой прикладных информационных технологий Кузбасского государственного технического университета им. Т.Ф. Горбачева, дали положительные отзывы на диссертацию.

Ведущая организация — ФГБОУ ВО «Новосибирский государственный технический университет» (НГТУ), в своем положительном заключении, рассмотренном на совместном научном семинаре кафедр теоретической и прикладной информатики и автоматизированных систем управления, составленном председателем научного семинара д.т.н. проф. Авдеенко Т.В. (протокол № 1 от 23.11.2020 г.) и утвержденном проректором по научной работе д.т.н., доц. Бровановым С.В., указала, что диссертационная работа Грибкова Е.И. является законченной научно-квалификационной работой, в которой изложены новые научно обоснованные технологические решения по извлечению структурированной информации из текстов пользователей о продуктах, имеющие существенное значение для повышения скорости и качества реагирования компаний-производителей на запросы потребителей, возникающие на этапах эксплуатации и сопровождения продуктов. Работа имеет существенное значение как в части развития методов извлечения составных объектов из текстов на естественном языке с использованием моделей на основе систем переходов и нейронных сетей, так и в области применения данного типа моделей при решении практических задач: анализе текстов отзывов пользователей о продуктах с целью извлечения мнений о товарах и продавце; анализе текстов обращений пользователей на этапе эксплуатации и сопровождения программного продукта с целью выявления запросов на новый функционал, программных ошибок и оценок существующих функций. Автор диссертационной работы Грибков Егор Игоревич заслуживает присуждения учёной степени кандидата технических наук по специальности 05.13.17 «Теоретические основы информатики».

Соискатель имеет 10 опубликованных работ, в том числе по теме диссертации 4 работы, опубликованные в рецензируемых научных изданиях, рекомендованных ВАК. Общий объем — 5,15 п.л, авторский вклад — 4 п.л. Имеется 2 свидетельства о регистрации программ для ЭВМ.

Наиболее значимые работы:

1. Ехлаков, Ю. П. Модель извлечения пользовательских мнений о потребительских свойствах товара на основе рекуррентной нейронной сети / Ю. П. Ехлаков, **Е. И. Грибков** // Бизнес-информатика. — 2018. — Т. 46, № 4. — С. 7—16.

2. **Грибков, Е. И.** Нейросетевая модель на основе системы переходов для извлечения составных объектов и их атрибутов из текстов на естественном языке / Е. И. Грибков, Ю. П. Ехлаков // Доклады ТУСУР. — 2020. — Т. 23, № 1. — С. 47—52.

3. **Грибков, Е. И.** Нейросетевая модель на основе системы переходов для извлечения и анализа тональности пользовательских мнений / Е. И. Грибков, Ю. П. Ехлаков // Искусственный интеллект и принятие решений. — 2020. — № 1. — С. 99—110.

На диссертацию и автореферат поступило 8 положительных отзывов из следующих организаций: Нижегородский государственный технический университет им. Р.Е. Алексеева (Ломакина Л.С., д.т.н., профессор кафедры «Вычислительные системы и технологии»); Институт систем информатики им. А.П. Ершова Сибирского отделения РАН, г. Новосибирск (Загорулько Ю.А., к.т.н., зав. лабораторией искусственного интеллекта); Научно-исследовательский вычислительный центр МГУ имени М. В. Ломоносова (Лукашевич Н.В., д.т.н., ведущий научный сотрудник лаборатории анализа информационных ресурсов); Национальный исследовательский университет «Высшая школа экономики», г. Москва (Авдошин С.М., к.т.н., профессор, руководитель департамента программной инженерии); Алтайский государственный университет, г. Барнаул (Оскорбин Н.М., д.т.н., профессор, профессор каф. теоретической кибернетики и прикладной математики); Сибирский государственный университет телекоммуникаций и информатики, г. Новосибирск (Канев В.С., д.т.н., профессор, зав. каф. математического моделирования бизнес-процессов); Алтайский государственный технический университет им. И.И. Ползунова, г. Барнаул (Крючкова Е.Н., к.ф.-м.н., доцент, профессор кафедры прикладной математики); Национальный исследовательский «Томский государственный университет» (Замятин А.В., д.т.н., доцент, директор Института прикладной математики и компьютерных наук, заведующий кафедрой теоретических основ информатики).

В отзывах на диссертацию и автореферат указаны следующие основные замечания: в ходе экспериментального исследования предложенных моделей использовались наборы текстов только на одном языке – русском, в связи с этим возникают вопросы о возможности обработки текстов на других языках, в частности, о наличии зависимости качества получаемых моделью предсказаний от языка текстов и возможности обрабатывать тексты на нескольких языках одной моделью; автором исследованы варианты использования моделей LSTM и CNN для извлечения признаков из исходного текста, однако существуют и другие достаточно распространенные варианты сетей, например рекуррентные сети Элмана, управляемые рекуррентные блоки (GRU), XL, BERT, GPT-2, GPT-3, Transformer, нейронные сети на основе механизмов внимания, не приведено сравнение предложенной модели с другими моделями извлечения структурированной информации из текстов пользователей (не нейросетевыми); при адаптации архитектуры нейросетевой модели для решения конкретной задачи следовало бы описать семантику фрагментов и атрибутов в виде онтологии модели предметной области; архитектура программных систем представлена в нотации схем AWS, следовало бы использовать более общий и распространенный аналог, например UML; для получения векторных представлений слов при обучении моделей использован метод fastText, при этом по тексту работы нет пояснений, каким образом конкретно слова преобразуются в вектора, в чем отличие от других аналогичных алгоритмов, почему выбор сделан в его пользу; не объяснен выбор конкретных значений гиперпараметров при обучении моделей, следовало бы выделить отдельную валидационную выборку и продемонстрировать оптимальность выбора параметров на диаграммах зависимости точности модели при заданном значении гиперпараметров; тональность задается шкалой, состоящей из трех уровней: положительной (+), нейтральной (0), негативной (-), однако на практике намного чаще используется не просто вектор, задающий оцениваемую сущность, а степень выраженности положительного или отрицательного тона; неясно, насколько репрезентативными были экспериментальные данные и какие уровни доверительной вероятности были использованы при экспериментальной апробации алгоритмов.

Выбор официальных оппонентов обосновывается тем, что д.т.н., проф. Спицын В.Г., является известным специалистом в области глубокого обучения, нейросетевых методов распознавания и классификации текстовой информации, семантической сегментации изображений. Д.т.н., проф. Пимонов А.Г. является признанным специалистом в области нейросетевых моделей прогнозирования производственно-технологических и социально-экономических процессов, интеллектуальных систем принятия решений на основе алгоритмов нечеткого логического вывода.

Выбор ведущей организации обосновывается тем, что ФГБОУ ВО НГТУ имеет высококвалифицированных специалистов в области методов автоматического извлечения знаний из массивов текстовой информации с использованием онтологического подхода, применения нейросетевых моделей сегментации изображений в прикладных задачах компьютерного зрения, которые имеют достаточный объём публикаций по тематике диссертации в ведущих изданиях и способны определить и аргументировано обосновать научную и практическую ценность диссертационной работы Грибкова Е.И.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

– предложена нейросетевая модель на основе системы переходов для извлечения составных объектов и их атрибутов из текстов на естественном языке, отличающаяся возможностью одновременно предсказывать структуру объекта и значения его атрибутов с возможностью адаптации модели под конкретные задачи через задание множеств, описывающих смысловое содержание фрагментов и атрибутов;

– разработана нейросетевая модель для извлечения и анализа мнений из текстов пользовательских отзывов о потребительских свойствах товаров отличающаяся от известных моделей использованием подхода на основе системы переходов и позволяющая получить показатели качества извлечения мнений по критерию меры точности и полноты: при определении фрагментов – $F_1=0,795$, при определении взаимосвязей – $F_1=0,723$, при определении атрибутов – $F_1=0,631$, что превосходит показатели качества извлечения мнений из текстов отзывов, полученные на

нейросетевых моделях-аналогах;

– разработана нейросетевая модель для анализа запросов пользователей на этапе эксплуатации и сопровождения программного продукта, отличающаяся от известных моделей использованием подхода на основе системы переходов и позволяющая получить показатели качества извлечения информативных фраз по критерию меры точности и полноты: при определении фрагментов – $F_1=0,633$, при определении взаимосвязей – $F_1=0,693$, что превосходит показатели качества извлечения информативных фраз пользователей, полученные на нейросетевых моделях-аналогах.

Теоретическая значимость исследования обоснована тем, что в работе проведена модернизация нейросетевых моделей извлечения составных объектов из текстов отзывов пользователей о продукте на естественном языке посредством использования системы переходов для одновременного предсказания структуры объекта, значения его атрибутов и процедуры адаптации модели при решении конкретных задач через задание множеств, описывающих смысловое содержание фрагментов и атрибутов.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что разработаны и внедрены нейросетевые модели для извлечения и анализа мнений из текстов пользовательских отзывов о продуктах и для извлечения и анализа запросов пользователей на этапе эксплуатации и сопровождения программного продукта, реализованные в виде программных систем «Quiddi Semantics» и «Quiddi Support Analyst». Внедрение осуществлено в коммерческих продуктах компании ООО «ТомскСофт»; в ФГБОУ ВО «ТУСУР» при выполнении государственного задания Министерства науки и высшего образования РФ, проект FEWM-2020-0036, в учебном процессе кафедры автоматизации обработки информации при проведении занятий по дисциплинам «Интеллектуальные вычислительные системы», «Анализ больших данных», «Нейронные сети и их применение» магистров по направлению 09.04.04 — «Программная инженерия». Практическое использование результатов работы позволит: предприятиям-произ-

водителям продуктов проводить модификацию комплекса маркетинговых мероприятий с целью улучшения положения продукта на рынке; ИТ-компаниям автоматизировать работу служб технической поддержки пользователей, проводить развитие программного продукта в соответствии с предложениями пользователей; потенциальным покупателям продуктов в интернет-магазинах при совершении покупок сравнивать характеристики потребительских свойств товаров-альтернатив.

Оценка достоверности результатов исследования выявила, что идея применения нейросетевых моделей на основе систем переходов при решении задач извлечения структурированной информации из текстов пользователей на естественном языке базируется на обобщении научных работ по тематике диссертации; теория построена на корректном использовании математического аппарата нейронных сетей при разработке моделей, согласуется с опубликованными в периодической печати научными результатами по теме диссертации. Адекватность предложенных в работе моделей подтверждается результатами их внедрения в продукты компании ООО «ТомскСофт».

Личный вклад соискателя состоит в разработке нейросетевых моделей и их реализаций в виде программных систем, проведении экспериментальных исследований и внедрении результатов, в разработке классификаторов типов мнений о продуктах, постановках задач исследования, подготовке публикаций по основным результатам работы совместно с научным руководителем.

Диссертация Грибкова Е. И. на соискание ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики», является законченной научно-квалификационной работой, в которой изложены новые научно обоснованные технологические решения по извлечению структурированной информации из текстов пользователей о продуктах, имеющие существенное значение для повышения скорости и качества реагирования компаний-производителей на запросы потребителей, возникающие на этапах эксплуатации и сопровождения продуктов, что соответствует требованиям пп. 9–14 «Положения о порядке присуждения ученых степеней».

На заседании 24 декабря 2020 г. диссертационный совет принял решение при-
судить Грибкова Е. И. ученую степень кандидата технических наук.

При проведении тайного голосования диссертационный совет в количестве
13 человек, из них 4 доктора наук по специальности рассматриваемой диссертации,
участвовавших в заседании, из 19 человек, входящих в состав совета, прогосо-
вали: за — 13, против — 0, недействительных бюллетеней — 0.

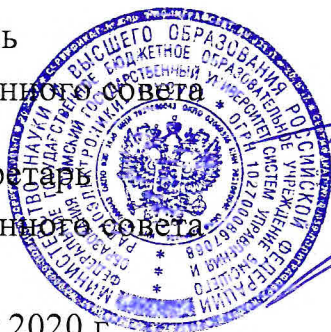
Председатель

диссертационного совета

Ученый секретарь

диссертационного совета

25» декабря 2020 г.



Ходашинский Илья Александрович

Костюченко Евгений Юрьевич