



УТВЕРЖДАЮ:

Проректор ТУСУРа

по научной работе и инновациям

к.т.н. доцент

А.Г. Лоцилов

« 08 » 10 2020 г.

## ЗАКЛЮЧЕНИЕ

федерального государственного бюджетного образовательного учреждения высшего образования «Томский государственный университет систем управления и радиоэлектроники»

Диссертация «Нейросетевые модели на основе системы переходов для извлечения структурированной информации о продуктах из текстов пользователей» выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Томский государственный университет систем управления и радиоэлектроники» (ТУСУР).

В 2015 году Грибков Егор Игоревич окончил ТУСУР по специальности «Автоматизированные системы обработки информации и управления», В 2017 году магистратуру ТУСУРа по направлению подготовки «Программная инженерия». В период подготовки диссертации соискатель Грибков Егор Игоревич обучался в очной аспирантуре ТУСУРа на кафедре автоматизации обработки информации (АОИ). Удостоверение о сдаче кандидатских экзаменов выдано в 2020 году ТУСУРом.

Научный руководитель — Ехлаков Юрий Поликарпович, доктор технических наук, профессор, профессор кафедры автоматизации обработки информации (АОИ) ТУСУРа.

### **По итогам обсуждения принято следующее заключение:**

Представленная диссертационная работа Грибкова Е.И. посвящена развитию методов извлечения структурированной информации из пользовательских текстов на естественном языке с использованием нейронных сетей в части извлечения из текстов и анализа составных объектов и их атрибутов.

В работе раскрыта актуальность темы исследования, произведена оценка степени ее проработанности как на теоретическом уровне, так и в практической плоскости. В ходе исследования проанализированы существующие подходы к анализу пользовательских текстов, выявлена проблема распространения ошибки между компонентами моделей, применяемых для извлечения структурированной информации из текстов. Для решения данной проблемы предлагается оригинальная архитектура нейросетевой модели на основе системы переходов



для извлечения составных объектов и их атрибутов из текстов. Предложенная архитектура адаптирована для решения двух прикладных задач: извлечения и анализа мнений из текстов пользовательских отзывов о продуктах; анализа запросов пользователей на этапе эксплуатации и сопровождения программного продукта. Проведено экспериментальное исследование моделей, показывающее превосходство предложенных моделей над рассмотренными альтернативными решениями. Представлено описание программных систем, реализующих предложенные модели и алгоритмы.

**Личный вклад автора** состоит в исследовании предметной области, теоретическом и экспериментальном исследовании разработанных моделей и алгоритмов, проектировании и реализации подсистем обучения моделей и анализа текстов в составе программных систем «Quiddi Semantics» и «Quiddi Support Analyst». Содержательные и математические постановки задач, структуры классификаторов типов мнений пользователей и типов информативных фраз о программном обеспечении были разработаны автором совместно с научным руководителем. Реализация программных систем осуществлена совместно с Трошиным М. и Пекарских Е.

**Достоверность** полученных результатов обусловлена корректным применением аппарата нейронных сетей при разработке моделей, что подтверждается публикациями результатов в профильных рецензируемых журналах, рекомендованных ВАК, результатами экспериментальных исследований предложенных моделей и аналогов, а также их практическим использованием в коммерческих программных продуктах компании ООО «ТомскСофт».

**Научная новизна** полученных в диссертационной работе результатов заключается в следующем.

1. Предложена оригинальная архитектура нейросетевой модели на основе системы переходов для извлечения составных объектов и их атрибутов из текстов на естественном языке, позволяющая одновременно предсказывать структуру объекта и значения его атрибутов, с возможностью адаптации под конкретные задачи через задание множеств, описывающих семантику фрагментов и атрибутов.

2. На основе предложенной архитектуры разработана оригинальная нейросетевая модель для извлечения и анализа мнений из текстов пользовательских отзывов о продуктах, отличающаяся от известных моделей использованием подхода на основе системы переходов и позволяющая получить лучшие показатели качества извлечения частей составных объектов: 0,795  $F_1$  при определении фрагментов; 0,723  $F_1$  при определении отношений; 0,631  $F_1$  при определении атрибутов, что на 2,38%, 1,40% и 17,40% соответственно превосходит показатели наилучшего альтернативного решения.

3. На основе предложенной архитектуры разработана оригинальная 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 лучшие показатели качества определения частей составных объектов: 0,633  $F_1$  при определении фрагментов; 0,693  $F_1$  при определении отношений,



что на 5,40% и 52,31% соответственно превосходит показатели наилучшего альтернативного решения.

**Практическая значимость.** Полученные в ходе написания диссертации результаты используются в коммерческих программных продуктах компании ООО «ТомскСофт» при реализации сервиса «Quiddi.ru» для оказания услуг населению о качестве продуктов. Разработанные нейросетевые модели, реализованные в виде программных продуктов «Quiddi Semantics» и «Quiddi Support Analyst», могут использоваться соответственно маркетологами при анализе текстов отзывов пользователей о продуктах, специалистами службы технической поддержки пользователей ИТ-компаний для обеспечения эксплуатации программного продукта в соответствии с его техническими характеристиками и развития продукта в соответствии с предложениями пользователей.

Ценность рассматриваемых научных достижений автора состоит в использовании полученных в диссертационной работе результатов при выполнении ТУСУРом государственного задания Министерства науки и высшего образования РФ, проект FEWM-2020-0036 «Методологическое и инструментальное обеспечение принятия решений в задачах управления социально-экономическими системами и процессами в гетерогенной информационной среде».

Полученные результаты используются также в учебном процессе кафедры автоматизации обработки информации при организации занятий по дисциплинам «Интеллектуальные вычислительные системы», «Анализ больших данных», «Нейронные сети и их применение» при подготовке магистров по направлению 09.04.04 «Программная инженерия».

По материалам диссертации Грибкова Е.И. опубликовано 10 работ, в том числе 4 статьи в журналах, включенных в Перечень российских рецензируемых научных журналов. В опубликованных работах достаточно полно отражены материалы диссертационного исследования.

#### *В рецензируемых журналах из списка ВАК*

1. 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

2. 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.

3. 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.

4. Грибков, Е.И. Нейросетевая модель для обработки запросов пользователей на этапе эксплуатации и сопровождения программного продукта / Е.И. Грибков, Ю.П. Ехлаков // Бизнес-информатика. — 2020. — Т. 14, № 1. — С. 7—18.



*Свидетельства о государственной регистрации программы для ЭВМ*

1. «Quiddi Semantics» / Е.И. Грибков [и др.] // Свидетельство о государственной регистрации программы для ЭВМ №2019612276 от 14.02.2019.
2. «Quiddi Support Analyst» / Е.И. Грибков [и др.] // Свидетельство о государственной регистрации программы для ЭВМ №2020614799 от 24.04.2020.

*Статьи и труды научно-практических конференций сборников*

1. Грибков, Е.И. Набор данных и модель глубокого обучения для анализа текстов отзывов пользователей / Е.И. Грибков, Ю.П. Ехлаков // Наука. Технологии. Инновации. Сборник научных трудов. В 9-ти частях. — Новосибирск: НГТУ, 2018. — С. 180—184.
2. Грибков, Е.И. Модель обработки запросов пользователей на этапе эксплуатации и сопровождения программного продукта / Е.И. Грибков, Ю.П. Ехлаков // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. — Томск: В-Спектр, 2019. — С. 141—143.
3. Грибков, Е.И. Модель извлечения структурированных объектов и их атрибутов из текстов на естественном языке / Е.И. Грибков // Сборник избранных статей научной сессии ТУСУРа (Томск, 22–24 мая 2019 г.): в 2 ч. — Томск: В-Спектр, 2019. — С. 54—56.
4. Грибков, Е.И. Модель на основе системы переходов для извлечения составных объектов из текстов / Е.И. Грибков, Ю.П. Ехлаков // Сборник избранных статей научной сессии ТУСУР, Томск, 13–30 мая 2020. — Томск: В-Спектр, 2020. — С. 52—55.

Диссертационная работа Грибкова Е.И. на тему: «Нейросетевые модели на основе системы переходов для извлечения структурированной информации о продуктах из текстов пользователей» соответствует требованиям «Положения о порядке присуждения ученых степеней», предъявляемым к кандидатским диссертациям, является законченным самостоятельным научным исследованием, имеющим научную ценность. Содержание диссертации соответствует паспорту специальности 05.13.17 – «Теоретические основы информатики» (технические науки): п. 5 «Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях, разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений»; п.6 «Разработка методов, языков и моделей человеко-машинного общения; разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения данных из текстов на естественном языке».

Диссертация «Нейросетевые модели на основе системы переходов для извлечения структурированной информации о продуктах из текстов пользователей» Грибкова Егора Игоревича рекомендуется к защите на соискание учёной степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики».

Заключение принято на заседании кафедры автоматизации обработки информации Томского государственного университета систем управления и радиоэлектроники.

Присутствовало на заседании 26 чел. Результаты голосования: «за» – 26 чел., «против» – 0, «воздержалось» – 0, протокол № 5 от 07 октября 2020 г.

Зав. кафедрой АОИ ТУСУРа  
кандидат экономических наук доцент

 А.А. Сидоров

Доцент кафедры АОИ ТУСУРа  
кандидат технических наук доцент

 Т.О. Перемитина