

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.087.01
на базе Федерального государственного бюджетного учреждения науки
Институт системного программирования Российской академии наук
Федерального агентства научных организаций РФ
по диссертации на соискание ученой степени кандидата наук

аттестационное дело № _____

решение диссертационного совета от 16 июня 2016 года № 2016/12

О присуждении Тутубалиной Елене Викторовне, гражданке РФ ученой степени кандидата физико-математических наук.

Диссертация «Методы извлечения и резюмирования критических отзывов пользователей о продукции» по специальности 05.13.11 – «математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» принята к защите 14 апреля 2016 года, протокол № 2016/09 диссертационным советом Д 002.087.01 на базе Федерального государственного бюджетного учреждения науки Институт системного программирования Российской академии наук (ведомственная принадлежность – Федеральное агентство научных организаций, адрес: 109004, г. Москва, ул. А. Солженицына, дом 25), создан Приказом Минобрнауки России о советах по защите докторских и кандидатских диссертаций от 2 ноября 2012 г. № 714/нк.

Соискатель Тутубалина Елена Викторовна, 1990 года рождения, работает в должности младшего научного сотрудника в федеральном государственном автономном образовательном учреждении высшего образования «Казанский (Приволжский) федеральный университет».

В 2012 году соискатель окончила федеральное государственное автономное образовательное учреждение высшего образования «Казанский (Приволжский) федеральный университет». В 2015 году соискатель окончила аспирантуру федерального государственного автономного образовательного учреждения высшего образования «Казанский (Приволжский) федеральный университет».

Диссертация выполнена в Высшей школе информационных технологий и информационных систем федерального государственного автономного учреждения высшего образования «Казанский (Приволжский) федеральный университет».

Научный руководитель – доктор физико-математических наук Соловьев Валерий Дмитриевич, ведущий научный сотрудник Научно-образовательного центра по лингвистике им. И.А. Бодуэна де Куртенэ федерального государственного автономного учреждения высшего образования «Казанский (Приволжский) федеральный университет».

Официальные оппоненты:

1. Ильин Вячеслав Анатольевич, доктор физико-математических наук, начальник Отдела информационных технологий и математического моделирования Курчатовского комплекса НБИКС-технологий НИЦ «Курчатовский институт».
2. Поляков Владимир Николаевич, кандидат технических наук, доцент кафедры автоматизированных систем управления федерального государственного автономного учреждения высшего образования «Национальный исследовательский технологический университет «МИСиС».

дали положительные отзывы на диссертацию.

Ведущая организация Федеральное государственное бюджетное учреждение науки Санкт-Петербургское отделение Математического института им. В.А. Стеклова (г. Санкт-Петербург) в своем положительном заключении, подписанном Матиясевицем Юрием Владимировичем, академиком РАН, доктором физико-математических наук, заведующим лабораторией математической логики ПОМИ РАН, указала, что диссертационная работа содержит новые научные результаты, имеющие существенное значение для науки и практики.

Выбор официальных оппонентов и ведущей организации обосновывается их компетентностью и достижениями в области интеллектуального анализа

данных, наличием публикаций в сфере исследований, соответствующей теме диссертации.

Соискатель имеет 13 опубликованных работ, в том числе по теме диссертации 10 работ, из которых 8 опубликованы в рецензируемых научных изданиях, в том числе 2 работы опубликованы в изданиях, входящих в список ВАК.

Наиболее значимые работы по теме диссертации:

1. Тутубалина Е. В. Совместная вероятностная тематическая модель для идентификации проблемных высказываний, связанных с нарушением функциональности продуктов //Труды Института системного программирования РАН. — 2015. — Т. 4, № 27. — С. 100—120.
2. Тутубалина Е. В. Извлечение проблем, связанных с неисправностями и нарушением функциональности продуктов, на основании отзывов пользователей // “Вестник КГТУ им. А. Н. Туполева”. — 2015. — Т. 3. — С. 139—146.
3. Ivanov V., Tutubalina E. Clause-based approach to extracting problem phrases from user reviews of products // Analysis of Images, Social Networks and Texts. — Springer International Publishing, 2014. — С. 229—236
Автору принадлежат введение, предложенный метод, экспериментальное исследование, выводы, заключение.
4. Tutubalina E. Dependency-Based Problem Phrase Extraction from User Reviews of Products // Text, Speech, and Dialogue. — Springer International Publishing, 2015. — С. 199—206.
5. Tutubalina E., Nikolenko S. Inferring Sentiment-Based Priors in Topic Models // Advances in Artificial Intelligence and Its Applications. — Springer International Publishing, 2015. — С. 92—104.
Автору принадлежат введение, обзор предметной области, предложенный алгоритм, экспериментальное исследование, выводы, заключение.
6. Tutubalina E. Target-Based Topic Model for Problem Phrase Extraction // Advances in Information Retrieval. — Springer International Publishing, 2015. — С. 271—277.

Диссертационный совет отмечает, что соискателем получены основные результаты:

- разработан метод классификации предложений, основанный на знаниях в виде созданных словарей и на контекстно-свободной грамматике, отражающей структуру сложных предложений;
- разработаны математическая модель и метод классификации предложений отзывов пользователей по отношению к объектам, связанным с предметной областью, на основе синтаксических связей слов и мер семантической близости;
- разработаны модели, учитывающие специфику отдельного отзыва и особенности коллекции критических отзывов предметной области в целом для выделения тематически сгруппированных объектов мнений;
- разработан программный комплекс для извлечения критических высказываний пользователей и построения тематических моделей.

Проведение экспериментальных исследований с помощью разработанного автором программного обеспечения подтвердило улучшение качества работы разработанных моделей и методов по сравнению с существующими методами.

Теоретическая значимость исследования состоит в том, что:

- разработаны новые алгоритмы выделения и анализа критических высказываний, основанные на построенной в диссертации контекстно-свободной грамматике проблемных высказываний;
- разработаны две вероятностные тематические модели, основанные на мультиномиальных распределениях во множестве слов, и получены формулы апостериорного распределения для определения оптимальных значений скрытых переменных в моделях.

Значение полученных соискателем результатов исследования для практики состоит в том, что:

- разработанные вероятностные модели и автоматические методы позволяют повысить качество классификации высказываний и тематического моделирования отзывов по сравнению с существующими моделями и методами;
- разработанные математические модели и методы извлечения высказываний могут быть использованы при решении прикладных задач анализа мнений: классификации текстовых документов, извлечения информации, кластеризации информации на основе тематических моделей;
- на базе разработанных методов создан программный комплекс для извлечения высказываний пользователей и построения тематических моделей, выложенный в открытый доступ, который может быть использован для решения аналогичных задач в других областях.

Достоверность результатов исследования подтверждается тем, что:

- формулы выведены на основе формальных математических преобразований;
- разработанные модели и методы оцениваются с помощью стандартных метрик качества систем анализа текстов на естественном языке;
- экспериментальные результаты классификации высказываний существующими методами машинного обучения с учителем получены с помощью программных средств, размещенных в открытом доступе;
- показана воспроизводимость результатов исследования с помощью разработанных моделей и методов на текстах различной длины на русском и английском языках о широкой группе товаров;
- выводы теории согласуются с полученными экспериментальными данными. Установлена согласованность результатов диссертации с результатами, представленными в независимых источниках в рамках задач анализа мнений, в тех случаях, когда сравнение допустимо.

Построенные в диссертации модели и алгоритмы методов извлечения и резюмирования критических отзывов пользователей соответствует пункту

«Модели, методы и алгоритмы проектирования и анализа программ и программных систем, их эквивалентных преобразований, верификации и тестирования» паспорта специальности 05.13.11. Разработанный программный комплекс соответствует пункту «Человеко-машинные интерфейсы; модели, методы, алгоритмы и программные средства машинной графики, визуализации, обработки изображений, систем виртуальной реальности, мультимедийного общения» паспорта специальности 05.13.11. Приведенные оценки качества работы систем анализа текстов на естественном языке относятся к пункту «Оценка качества, стандартизация и сопровождение программных систем» паспорта специальности 05.13.11.

Личный вклад соискателя состоит в проведении исследования предметной области, формулировании научных положений и выводов, сборе исходных данных, разработке предложенных методов и вероятностных моделей, проведении экспериментальных оценок, интерпретации полученных результатов, программной реализации предложенных алгоритмов, подготовке основных публикаций и участии в апробации результатов.

На заседании 16 июня 2016 года диссертационный совет принял решение присудить Тутубалиной Елене Викторовне ученую степень кандидата физико-математических наук.

При проведении тайного голосования диссертационный совет в количестве 15 человек, из них 8 докторов наук по специальности рассматриваемой диссертации, участвовавших в заседании, из 20 человек, входящих в состав совета, проголосовали: за – 14, против – 1, недействительных бюллетеней – 0.

Председатель диссертационного совета,
академик РАН

Иванников В.П.

Ученый секретарь диссертационного совета,
кандидат физико-математических наук

Зеленов С.В.

16 июня 2016 года.