

ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

доктора физико-математических работ Соловьева Валерия Дмитриевича

на диссертационную работу Тутубалиной Елены Викторовны «Методы извлечения и резюмирования критических отзывов пользователей о продукции», представленную к защите на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» в Диссертационный совет Д 002.087.01 при Институте системного программирования Российской академии наук

Последнее десятилетие характеризуется развитием технологий в сфере автоматической обработки текстов на естественном языке, публикуемых на страницах различных web-ресурсах таких, как социальные сети, сайты интернет-сервисов. На рынке потребительских товаров наблюдается резкая динамика увеличения количества технически сложных товаров, поэтому одной из актуальных и важных задач является задача анализа мнений авторов текстов по отношению к продуктам, обсуждаемых в тексте и связанных с претензиями авторов по поводу удобства использования наряду с ненадлежащим качеством.

Однако в настоящее время задача автоматического извлечения высказываний, связанных с неисправностями и нарушением функциональности продуктов, выполняется, как правило, лишь с помощью лингвистических правил на основе ключевых слов, базовых тематических моделей и методов машинного обучения на небольшом наборе признаков.

Диссертационная работа Тутубалиной Е.В. посвящена разработке методов извлечения информации о высказываниях пользователей, содержащих указания на трудности в использовании продуктов. Основными требованиями к предлагаемым методам являются более высокое качество по сравнению с существующими методами и переносимость методов на тексты различных языков о широкой группе товаров.

В своей диссертационной работе Елене Викторовне удалось предложить метод классификации предложений, основанный на знаниях в виде созданных словарей и правилах, учитывающий грамматическую структуру сложных предложений относительно союзов. Предложенный метод улучшил качество классификации текстов о продуктах пяти предметной областей на русском и английских языках по сравнению с существующими методами машинного обучения с учителем. Кроме того, предложен метод классификации предложений отзывов по отношению к целевым объектам, связанных с предметной областью. Метод, основанный на использовании синтаксических связей слов в предложении и мер семантической связанности,

показывал улучшение качества классификации текстов о механических продуктах.

Тутубалиной Е.В. удалось предложить две вероятностные модели для выделения тематически сгруппированных объектов мнений, указывающих на проблемные ситуации. Предложенные модели, основанные на методе латентного размещения Дирихле и учитывающие новые скрытые переменные, улучшили несколько критериев качества, предъявляемых к такого рода моделям, по сравнению с существующими вероятностными моделями. Кроме того, качественный анализ результатов подтвердил улучшение интерпретируемости тем, которые отражают различия тематических распределений применительно к различным претензиям авторов об удобстве использования или технических неполадках продуктов.

За время работы над диссертацией Тутубалина Е.В. проявила себя как компетентный специалист в области обработки текстов на естественном языке. В ходе диссертационного исследования Елена Викторовна показала владение современными методами обработки текстов на естественном языке и математическим аппаратом тематических вероятностных моделей. Отдельно следует отметить самостоятельность проделанной научно-исследовательской работы, умение осуществлять экспериментальную проверку сформулированных гипотез и выполнять программную реализацию предложенных алгоритмов. Все результаты, представленные в диссертации, являются новыми. Программы, созданные в рамках диссертационного исследования, выложены в открытый доступ.

Считаю, что представленная диссертационная работа соответствует всем требованиям, предъявляемым Высшей аттестационной комиссией при Министерстве образования и науки Российской Федерации к работам на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей», а ее автор, Тутубалина Елена Викторовна, заслуживает присуждения ей ученой степени кандидата физико-математических наук.

Научный руководитель,
ведущий научный сотрудник Научно-образовательного центра по лингвистике им. И.А. Бодуэна де Куртенэ Института филологии и межкультурной коммуникации КФУ,
доктор физико-математических наук

Соловьев В.Д.

«21» марта 2016 года.