

ОТЗЫВ

об автореферате диссертации на соискание ученой степени кандидата филологических наук Марины Валерьевны Чернышевич на тему «Система автоматического сентимент-анализа текстов на английском языке», представленной к защите по специальности

10.02.21 – Прикладная и математическая лингвистика

В настоящее время в научной сфере и в сфере промышленных разработок по интеллектуальному анализу текстовых данных необходимы комплексные инструменты, ориентированные на обработку корпусов текстов больших объемов, извлекаемых автоматически из веб-ресурсов (электронные СМИ, социальные сети, электронные библиотеки и т.д.). К числу первоочередных задач, которые решаются при лингвистическом анализе таких ресурсов, относится автоматический сентимент-анализ текста (АСАТ), реализуемый как многоэтапная процедура. В АСАТ входит распознавание сентиментов – тональных слов и конструкций, выявление аспектов тональности, тональная классификация текстов и многое другое. С одной стороны, выполнение АСАТ на высоком уровне предполагает морфо-синтаксический и семантический анализ обрабатываемого текста. С другой стороны, сама процедура АСАТ может выполняться в разных объемах и с разной степенью детализации параметров входного корпуса текстов. Решение, предлагаемое в диссертации М.В.Чернышевич, удовлетворяет максимальному набору требований, предъявляемых к современным системам сентимент-анализа, что определяет актуальность проведенного исследования.

Разработка М.В.Чернышевич опирается на состоятельную теоретическую платформу, включающую многоуровневое лингвистическое описание текста, и на гибридный комплекс алгоритмов, объединяющий машинное обучение и систему языковых правил, эффективность которого доказана многими исследованиями автора. В диссертации обсуждается АСАТ как внедренный высокоэффективный исследовательский продукт, используемый в составе промышленного многоязычного лингвистического процессора системы инженерии и управления знаниями Goldfire Innovator. Это свидетельствует о высокой теоретической значимости проведенного исследования и реальном практическом применении его результатов.

Новаторский характер исследования, выполненного М.В.Чернышевич, определяется прежде всего тем, что в нем реализована расширенная шкала тональности, учитывающая многообразные потребности пользователей системы АСАТ, а также тем, что сам сентимент-анализ комбинируется с очень важными сопутствующими процедурами, прежде всего, с фильтрацией спам-сообщений и текстов с недостоверной оценкой, что в итоге приводит к качественно высокому результату автоматической обработки текстов и высоким оценкам функциональности системы в целом (для контрольной выборки точность составляет 86,09%, полнота составляет 81,80%). Основные исследовательские результаты представлены в 9 публикациях автора, 5 из них изданы в рецензируемых изданиях.

Автореферат диссертации является информативным и дает четкое представление о проведенной исследовательской работе. Текст автореферата составлен в соответствии с требованиями, предъявляемыми к данному типу научных сочинений.

По ходу знакомства с авторефератом у рецензента возникли вопросы, носящие дискуссионный характер и не подвергающиеся сомнению высокий уровень диссертационного исследования. Лингвистические данные, на материале которых проведено исследование, включают в себя разножанровые тексты на английском языке (сообщения из социальных сетей, новостные статьи, научные тексты).

1. Какие процедуры проводятся для расширения функционала действующей системы АСАТ с тем, чтобы она могла обрабатывать тексты на русском, белорусском и других языках?

2. Требуется ли перестройка шкалы тональности при обработке текстов на других языках?

3. Насколько предлагаемая шкала тональности чувствительна к жанрово-стилевой специфике обрабатываемых текстов (есть ли зоны этой шкалы, например, более частотные при анализе текстов социальных сетей и менее частотные при обработке научных текстов)?

Диссертационное исследование М.В.Чернышевич на тему «Система автоматического сентимент-анализа текстов на английском языке» носит новаторский характер, выполнено на высоком научном уровне, является оригинальным по замыслу и самостоятельным с точки зрения его реализации, тем самым, оно полностью соответствует требованиям, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор заслуживает присуждения искомой ученой степени.

Рецензент:

кандидат филологических наук (специальность 10.02.21), доцент,
доцент кафедры математической лингвистики
филологического факультета СПбГУ
Митрофанова Ольга Александровна

Рабочий адрес: Россия, 199034,
Университетская набережная,
дом 11, аудитория 193

E-mail: o.mitrofanova@spbu.ru

Мобильный телефон: 8(911)718-12-16

Личную подпись заверяю *Митрофанова О. А.*
Документ подготовлен по личной
инициативе

19 ДЕК 2019

Текст документа размещен в открытом
доступе на сайте СПбГУ по адресу
<http://spbu.ru/science/expert.html>

Ведущий специалист по ИИТ
Семезов В. В. Семезов

