

ОТЗЫВ

о диссертации Куликова Сергея Юрьевича
«Автоматическое извлечение мнений: лингвистический аспект»,
представленной на соискание ученой степени
кандидата филологических наук
по специальности 10.02.21
Прикладная и математическая лингвистика

Диссертационная работа С.Ю. Куликова посвящена исследованию активно развивающегося направления в современной прикладной лингвистике – автоматическому извлечению мнений. Этап анализа мнений, как правило, является надстройкой над базовыми лингвистическими сервисами, и качество его работы во многом зависит от настроек и точности нижележащих модулей, а так же специализированных лингвистических ресурсов. С практической точки зрения это приводит к особым сложностям при проектировании такого рода систем. С другой стороны, несмотря на огромное количество публикаций в данной области, почти отсутствует общепринятая теоретическая база. Вследствие вышесказанного **актуальность такого рода исследований несомненна.**

Прежде всего, хотелось бы отметить устранение Сергеем Юрьевичем терминологического изъяна в данной области: он не только разделяет задачи sentiment анализа и автоматического извлечения мнений (sentiment analysis and opinion mining), но и систематизирует подзадачи в каждом из направлений.

Заслуживает внимания скрупулезная работа над лингвистическими аспектами в исследуемой предметной области, поскольку зачастую это игнорируется в современных методах машинного обучения.

Немаловажным является и подробный анализ существующих на рынке систем, методов и подходов к решению задач автоматического извлечения мнений (первая часть диссертации), поскольку на сегодняшний день нет четко понимания всех возможностей и необходимости их применения в разных областях человеческой деятельности. Особенно важным является обоснование оценочных критериев результатов работы данных систем.

Подробное описание структуры систем автоматического извлечения мнений, основанных прежде всего на лингвистических предпосылках (представленное во второй части диссертации), не вызывает нареканий. Однако следует отметить, что недостаточное внимание уделено нейросетевым подходам в данной области. Несмотря на то, что на

сегодняшний день нейросети и глубокое машинное обучение не демонстрирует высокого качества результата в области лингвистической обработки текста, тем не менее, оно имеет явные перспективы. Особенно интересны гибридные модели, где нейросетевые технологии используются для оптимизации промежуточных результатов. Самым простым примером является использование Байесовских или Марковских сетей для реализации стохастических лингво-правил sentiment анализа. Это позволяет отразить вероятностные стороны языковой коммуникации и существенно улучшить результаты.

Некоторые пожелания вызывает третья часть исследования, посвященная специализированным ресурсам для извлечения и анализа мнений. В частности, автоматизированный сбор словника оценочных лексем. Автор отмечает, что список зависит от предметной области, а в основе оценки лексемы, как правило, лежит ее частотность. Вследствие этого почти все современные системы извлечения и анализа мнений имеют либо однополярную (с назначенным порогом срабатывания), либо двуполярную (позитив-негатив) шкалу оценки эмоционального компонента высказывания. Однако есть ряд ресурсов, где осуществлена попытка построения мультимодальных систем анализа тональности и анализа мнений. Это позволяет выявить ряд важных факторов, отражающих эмоциональные аспекты коммуникативной деятельности как агрессивность, депрессию, возбужденность и т.д., что, в свою очередь, может послужить материалом для систем прогнозирования. Такой подход требует не только создание узко специализированных ресурсов, сколько разработку единой теории лингвистических эмоций, основанной на эмотивной составляющей слов, которая развивается в процессе коммуникативной деятельности. И хотя, и автор сообщает, что в своем исследовании данной задачей не занимается (стр.28 п.1.7.3.), все-таки хотелось бы пожелать ему развивать теорию анализа субъективных высказываний в направлении исследования эмотивности лексики и разработки критериев мультимодальной оценки текста.

Диссертационное исследование С.Ю. Куликова обладает несомненной *научной новизной, теоретической и практической значимостью. Достоверность* работы обеспечивается проверкой представленных в работе результатов в промышленных системах лингвистической обработки текста.

Основные результаты работы были представлены на международных, всероссийских и российских конференциях, отражены в 24 публикациях (в т.ч. в изданиях, рекомендованных ВАК).

Все изложенное позволяет утверждать, что работа Куликова Сергея Юрьевича отвечает требованиям, предъявляемым на соискание ученой степени кандидата филологических наук по специальности 10.02.21 – Прикладная и математическая лингвистика, а ее автор заслуживает тем самым искомой степени.

к. ф. н., лингвист-программист ООО «2толк»

Соловьев Алексей Николаевич

