

УДК 004.8

ИССЛЕДОВАНИЕ ПОДХОДОВ РЕШЕНИЯ ЗАДАЧИ АСПЕКТНОГО АНАЛИЗА ТОНАЛЬНОСТИ ТЕКСТОВЫХ СООБЩЕНИЙ

Абрегова З.Х.(Университет ИТМО)

Научный руководитель – кандидат технических наук, доцент Махныткина О.В.
(Университет ИТМО)

Введение. Исследования выполнены за счет финансирования университета ИТМО в рамках НИР № 623088 «Разработка русскоязычного персонифицированного эмоционального диалогового агента». Исследование подходов к решению задачи аспектного анализа тональности текстовых сообщений имеет особое значение при разработке эмоционального диалогового агента, так как в одной реплике пользователь может высказывать различное мнение о нескольких объектах, процессах. Решение данной задачи требует глубокого понимания семантической структуры текста, а также учета контекста и субъективности высказываний [1].

Основная часть. Анализ тональности определяет эмоциональную окраску сообщения (позитивную, негативную или нейтральную). Отличие аспектного анализа тональности от классического анализа тональности заключается в выделении из текстового сообщения отдельных аспектов. Аспект - это характеристика или атрибут объекта, которая может вызвать проявление эмоциональной реакции у человека [2]. Вместо общей оценки текстового сообщения осуществляется оценка тональности отдельных фрагментов, предложений или аспектов, что может привести к более точным результатам оценки. Существует несколько методов анализа тональности: лингвистический подход, машинное обучение и гибридный подход [3]. Один из самых распространенных подходов к построению моделей аспектного анализа – это использование алгоритмов машинного обучения. Лучшие результаты для задачи классификации показывают BERT-based модели. Модели Rubert-tiny2 и RuELECTRA-small хорошо работают с русским текстом. Использование ансамбля данных моделей позволило получить более точные результаты при работе с текстами на русском языке. Для проведения экспериментов использовался корпус CABSAR (Corpus for Aspect-Based Sentiment Analysis in Russian-language). Корпус данных разделен на обучающую и тестовую выборки, содержит 5552 уникальных предложения (6705 пар аспект-сентимент).

Выводы. Проведен анализ подходов к решению задачи оценки тональности текстовых сообщений по выделенным аспектам. Предложено использование ансамбля моделей для Rubert-tiny2 и RuELECTRA-small для аспектного анализа тональности.

Список использованных источников:

1. Семина Т.А. Анализ тональности текста: современные подходы и существующие проблемы // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 6, Языкознание: Реферативный журнал. – 2020. №4. – С. 47 - 64.
2. Проноза, Е.В., Ягунова, Е.В. Аспектный Анализ Отзывов О Ресторанах Для Рекомендательных Систем Е-Туризма // Сборник научных статей XVIII Объединенной конференции "Интернет и современное общество". – Санкт-Петербург, 2015. – С. 130-141.
3. Котельников Е. В., Клековкина М. В. Автоматический анализ тональности текстов на основе методов машинного обучения // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Международной конференции «Диалог». – М: РГГУ, 2012. – С. 15–21.