

УДК 004.9

МЕХАНИЗМ ВЫЯВЛЕНИЯ ОНТОЛОГИЙ ПРИ ОПРЕДЕЛЕНИИ ОТНОШЕНИЯ ПОКУПАТЕЛЕЙ К ТОВАРУ ИЗ ОТЗЫВОВ НА МАРКЕТПЛЕЙСЕ

Груздева А.С. (Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

Юрьев Р.Н. (Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

Научный руководитель – д.т.н., проф. Бессмертный И.А.

(Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

Доклад посвящен этапу исследования новой квантово-подобной волновой модели представления текстовой информации и возможности ее применения для изучения общественного мнения.

Работа выполнена в рамках магистерско-аспирантской НИР № 620164 «Методы искусственного интеллекта для киберфизических систем».

Введение. В настоящее время постоянно растет интерес к анализу и исследованию общественного мнения в сферах торговли, производства, социологии и политики. С одной стороны, умножается число источников, содержащих текстовую информацию об отношении людей к товарам, услугам и событиям. Это социальные сети, специализированные сайты отзывов, маркетплейсы, которые дают возможность не только продавать товары, но и получать комментарии покупателей. С другой стороны растет понимание того, что для успешного функционирования в условиях динамично меняющейся ситуации, требуется эффективная обратная связь. Безусловно, в сфере изучения общественного мнения многие годы существуют и результативно применяются различные методы и алгоритмы, такие, как, например, сентимент – анализ. Тем не менее, разработка новых методик может расширить возможности существующих и дать новые перспективные результаты. Данная работа посвящена разработке и исследованию новой квантово-подобной модели представления текстовой информации и изучению возможностей ее применения для анализа общественного мнения.

Основная часть. В основе волновой модели лежит представление текста в виде ансамбля элементарных частиц. В соответствии с основами квантовой механики поведение такого ансамбля описывается при помощи набора сферических волн, а амплитуда вероятности обнаружения ансамбля частиц в некоторой точке пространства соответствует суммарной интенсивности волнового пакета в данной точке. В рамках волновой модели каждому слову текста, относящемуся к самостоятельной части речи, соответствует сферическая волна, обладающая своей амплитудой, волновым числом и начальной фазой. Пространство, в котором существуют данные волны, описывается при помощи дистрибутивно-семантической языковой модели, а расстояния, соответственно, являются величинами, обратно пропорциональными семантической близости. Такой подход позволяет вычислить амплитуду вероятности обнаружения ансамбля частиц (текста) в области близкой к термину, обозначающему некоторый класс и таким образом провести классификацию. На данном этапе работы для экспериментального исследования волновой модели была выбрана задача определения отношения покупателей к товару, исходя из отзывов на маркетплейсе. На сайте <https://www.wildberries.ru/> были отобраны комментарии, посвященные смартфонам, и произведено их распределение между классами «позитивный» и «негативный». Объективный класс определялся исходя из соответствующей оценки товара, поставленной автором отзыва: одна и две звезды расценивались, как негативный класс, четыре и пять – позитивный класс, три звезды определяют нейтральный отзыв, который не может быть с уверенностью отнесен ни к одному из классов. В результате проведенных расчетов квантово-

подобная модель показала большую точность по сравнению с классической методикой. Из недостатков волновой модели была отмечена более высокая трудоемкость, что влечет за собой увеличение количества времени, требуемого на выполнение классификации. Дальнейшие исследования будут направлены на повышение точности и снижение времени классификации с использованием волновой модели. Для этого планируется рассмотреть использование в качестве базиса других языковых моделей, а также уточнение алгоритмов вычисления параметров разрабатываемой модели.

Выводы. Экспериментальное исследование показало, что волновая модель может рассматриваться, как перспективная методика в сфере исследования общественного мнения, и может использоваться для решения задачи извлечения онтологий из неструктурированных текстовых данных. Данная модель обладает большей точностью, по сравнению с классическим подходом, однако требует дальнейшего развития и оптимизации.