

УДК 004.934

КОНЦЕПЦИЯ И АНАЛИЗ НАСТРОЕНИЙ В СОЦИАЛЬНЫХ СЕТЯХ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

БАРЩЕВСКИЙ Евгений Георгиевич

кандидат технических наук, профессор

Государственный университет морского и речного флота им. адмирала С.О. Макарова
г. Санкт-Петербург, Россия

Актуальность работы обусловлена широким использованием искусственного интеллекта как широкого спектра технологий, используемых для изучения и анализа настроений клиентов в социальных сетях.

Ключевые слова: искусственный интеллект, анализ настроений, анализ намерений, контекстный семантический поиск.

Анализ настроений – это контекстный анализ текста, который идентифицирует и извлекает субъективную информацию из исходных материалов и помогает бизнесу понять социальные настроения в отношении своего бренда, продукта или услуги при мониторинге онлайн-разговоров. Однако анализ потоков социальных сетей обычно ограничивается только базовым анализом настроений и показателями, основанными на подсчете. Это все равно, что просто царапать поверхность и упускать те ценные идеи, которые ждут своего открытия. Благодаря недавним достижениям в области глубокого обучения способность алгоритмов анализировать текст значительно улучшилась. Творческое использование передовых методов искусственного интеллекта может стать эффективным инструментом для проведения глубоких исследований. В настоящее время можно классифицировать входящие разговоры с клиентами о бренде по следующим признакам [4;6; 3]:

1. Ключевые аспекты продуктов и услуг бренда, которые волнуют клиентов.

2. Основные намерения и реакции пользователей относительно этих аспектов.

Если использовать сочетание базовых концепций, то это сочетание становится очень важным инструментом для анализа миллионов разговоров о бренде с точностью до человеческого уровня. Рассмотрим основные строительные блоки классификатора текста [1; 5].

Анализ настроений – это наиболее распро-

страненный инструмент классификации текста, который анализирует входящее сообщение и определяет, является ли лежащее в его основе настроение положительным, отрицательным или нейтральным.

Анализ намерений – инструмент, который анализируя намерения пользователей, стоящие за сообщением определяет, относится ли оно к мнению, новостям, маркетингу, жалобе, предложению, оценке или запросу.

Контекстный семантический поиск (CSS). Вот здесь все становится действительно интересно. Чтобы получить полезную информацию, важно понять, какой аспект бренда обсуждает пользователь. Например: Amazon хотела бы разделить сообщения, которые касались: задержек поставок, проблем с выставлением счетов, запросов, связанных с продвижением, обзоров продуктов и т. д. С другой стороны, Starbucks хотела бы классифицировать сообщения на основе того, связаны ли они с поведением персонала, новыми вкусами кофе, отзывами о гигиене, онлайн-заказами, названием и местоположением магазина и т. д. Для решения подобных проблем целесообразно использовать алгоритм интеллектуального поиска под названием «Контекстный семантический поиск» (также известный как CSS). Принцип работы CSS заключается в том, что он принимает тысячи сообщений и концепцию (например, Price) в качестве входных данных и фильтрует все сообщения, которые точно соответствуют данной концепции. Опыт показыва-

ет насколько CSS лучше по сравнению с существующими методами, используемыми в отрасли. Традиционный подход к фильтрации всех сообщений, связанных с ценой, заключается в поиске по ключевым словам «Цена» и другим тесно связанным словам, таким как (цена, плата, \$, оплачено). Однако этот метод не очень эффективен, поскольку практически невозможно придумать все релевантные ключевые слова и их варианты, которые представляют конкретную концепцию. CSS, с другой стороны, просто принимает имя концепции (Price) в качестве входных данных и фильтрует все контекстуально похожее, даже если очевидные варианты ключевого слова понятия не упоминаются. Техника искусственного интеллекта используется для преобразования каждого слова в определенную точку гиперпространства, а расстояние между этими точками используется для идентификации сообщений, контекст которых похож на исследуемую нами концепцию. Рассмотрим CSS в действии и то, как он работает с комментариями, связанными с Uber, на примерах ниже [1; 5; 2].

UBER: глубокий анализ.

Uber, самый ценный стартап в мире, стал пионером в области экономики совместного использования. Работая более чем в 500 городах по всему миру и обслуживая огромную базу пользователей, Uber получает множество отзывов, предложений и жалоб от пользователей. Часто социальные сети являются наиболее предпочтительным средством для регистрации таких проблем. Огромный объем входящих данных усложняет анализ, категоризацию и получение информации. Можно проанализировать онлайн-разговоры, происходящие в цифровых медиа, по нескольким темам продуктов: отмена, оплата, цена, безопасность и обслуживание. Для широкого охвата источников были взяты данные из последних комментариев на официальной странице Uber в Facebook, твитов с упоминанием Uber и последних новостных статей об Uber. Вот распределение точек данных по всем каналам:

1. Facebook: 34 173 комментариев.
2. Твиттер: 21 603 твита.
3. Новости: 4 245 статей.

Анализ настроений пользователей может дать представление об общем восприятии

бренда. Но чтобы копнуть глубже, важно дополнительно классифицировать данные с помощью контекстного семантического поиска. Если использовать алгоритм контекстного семантического поиска в том же наборе данных, принимая во внимание вышеупомянутые категории (Отмена, Оплата, Цена, Безопасность и Сервис).

ФЕЙСБУК.

Анализ настроений. Примечательно, что комментарии, относящиеся ко всем категориям, имеют в основном негативную направленность, за исключением первой. Количество положительных комментариев в отношении Прайса превысило количество отрицательных. Чтобы копнуть глубже, были проанализированы намерения этих комментариев. Facebook является социальной платформой, поэтому комментарии представляют собой множество *случайного контента, новостных репостов, маркетингового и рекламного контента, а также спама, ненужного и несвязанного контента*. Таким образом, если удалить все такие нерелевантные категории намерений и воспроизвести результат, то видно, что есть заметные изменения в настроениях, связанных с каждой категорией. Особенно в комментариях, связанных с ценами, где количество положительных комментариев упало с 46% до 29%. Это дает нам представление о том, как CSS может генерировать глубокую информацию из цифровых медиа. Таким образом, бренд может анализировать такие твиты и опираться на их положительные моменты или получать отзывы на отрицательные.

ТВИТТЕР.

Анализ настроений. Аналогичный анализ был проведен для просканированных твитов. При первоначальном анализе твиты, связанные с платежами и безопасностью, имели неоднозначную оценку. Анализ показал, что заметно сократилось количество положительных твитов, связанных с платежами. Кроме того, наблюдается значительное снижение количества положительных твитов по категории «Безопасность» (и связанным с ней ключевым словам). Кроме того, «Отмена», «Оплата» и «Услуга» (и связанные с ними слова) – самые обсуждаемые темы в комментариях в Твитте-

ре. Кажется, больше всего люди говорили о том, что водители отменяют поездку, и о взимаемой с них плате за отмену. Такой бренд, как Uber, может положиться на такую информацию и действовать по наиболее важным темам. Например, твиты, связанные с услугами, имели самый низкий процент положительных твитов и самый высокий процент негативных. Таким образом, Uber может анализировать такие твиты и реагировать на них, чтобы улучшить качество обслуживания. Понятно, что безопасность стала самой обсуждаемой темой в новостях. Интересно, что настроения в новостях в целом положительные, а также в каждой кате-

гории по отдельности.

Заключение. С развитием технологий наступила эпоха получения значимой информации из данных социальных сетей. Пример использования Uber дает вам представление о возможностях контекстного семантического поиска. В последнее время компании активно используют возможности данных, но чтобы получить самую глубокую информацию, необходимо использовать возможности искусственного интеллекта, глубокого обучения и интеллектуальных классификаторов, таких как контекстный семантический поиск и анализ настроений.

СПИСОК ЛИТЕРАТУРЫ

1. 13 удивительных способов использования технологии Emotion AI. – URL:<https://www.gartner.com/smarterwithgartner/13-surprising-uses-for-emotion-ai-technology/> (дата обращения: 03.04.2023).
2. Эмоции ИИ будут персонализировать взаимодействие. – URL:<https://www.gartner.com/smarterwithgartner/emotion-ai-will-ersonalize-interactions/> (дата обращения: 04.04.2023).
3. *Jakobi 'A. Proximity-Driven Motives in the Evolution of an Online Social Network // The Rise of Big Spatial Data.* Cham: Springer International Publishing, 2017. P. 197-209. DOI: 10.1007/978-3-319-45123-7_15.
4. *Mislove A. Pulse of the Nation: US Mood Throughout the Day Inferred from Twitter.* Northeastern University Khoury College of Computer Sciences, 2017. URL:<http://www.ccs.neu.edu/home/amislove/twittermood/> (дата обращения: 25.01.2022).
5. National Happiness Index Monitoring using Twitter for Bilanguages / D. Wang [и др.] // *Social Network Analysis and Mining.* 2021. Т. 11, № 1 P. 1-18. DOI: 10.1007/s13278-021-00728-0.
6. Social Data Analysis of Brazilian's Mood from Twitter / D.N. Prata [и др.] // *International Journal of Social Science and Humanity.* – 2016. – Т. 6, № 3. – P. 179-183. – DOI: 10.7763/IJSSH.2016.V6.640.

UDC 004. 934

CONCEPT AND ANALYSIS OF SENTIMENT IN SOCIAL NETWORKS USING ARTIFICIAL INTELLIGENCE

BARSHCHEVSKY Eugene Georgievich

Candidate of Sciences in Technology, Professor

Admiral Makarov State University of Maritime and Inland Shipping

St. Petersburg, Russia

The relevance of the work is due to the widespread use of artificial intelligence as a wide range of technologies used to study and analyze customer sentiment on social networks.

Keyword: artificial intelligence, sentiment analysis, intent analysis, contextual semantic search.