Елена Леонидовна Вартанова

Московский государственный университет им. М. В. Ломоносова

eva@smi.msu.ru

Мария Евгеньевна Аникина

Московский государственный университет им. М. В. Ломоносова

maria-anikina@yandex.ru

Юлия Сергеевна Нефедова

Московский государственный университет им. М. В. Ломоносова

yulja.nefedova@mail.ru

**О применении нейросетевой модели для анализа конфликтогенности в цифровой медиасреде**

Изучаются вопросы влияния репрезентации социального конфликта в СМИ на уровень агрессии аудитории. С помощью разработанной речевой нейросетевой модели проводится подробный анализ уровня конфликтогенности постов и комментариев к ним в социальной сети Telegram для разных субъектов медиапространства.

Ключевые слова: социальный конфликт, медиа, нейросетевая модель

Исследование выполнено за счет средств гранта Российского научного фонда (проект № 22-18-00225).

Современный уровень развития технологий искусственного интеллекта позволяет обрабатывать текстовую информацию большого объема в автоматическом режиме. В связи с этим становятся доступными различные исследования цифрового медиапространства, затрагивающие анализ большого числа сообщений его участников. В данной работе была разработана специальная математическая модель, позволяющая изучать влияние репрезентации социального конфликта в цифровых СМИ на уровень агрессии аудитории [1].

В качестве примера был рассмотрен арабо-израильский конфликт 2023 г., когда вооруженные силы ХАМАС вторглись на территорию Израиля. Исследования затронули 31 день указанного конфликта – с 7 октября по 6 ноября включительно. Исходными данными послужили посты и комментарии пользователей в социальной сети Telegram. С помощью рейтингов ресурса TGStat были отобраны каналы неинституционализированных СМИ и блогеров, позволяющие оставлять комментарии к своим постам. При этом общее число сообщений в коллекции превысило 300 тысяч.

На первом этапе в работе была составлена подробная карта событий арабо-израильского конфликта, а затем разработана математическая модель его репрезентации различными медиасубъектами. Модель показывает, что распределение числа комментариев по дням конфликта полностью соответствует распределению числа постов. При этом рассматриваемый период репрезентации с 7 октября по 6 ноября можно четко разделить на три фазы:

– 1 фаза (с 7 октября по 13 октября) характеризуется большим числом постов и комментариев в первые дни и заметным снижением числа публикаций в последующие;

– для 2 фазы (с 14 октября по 21 октября) характерна тенденция плавного снижения числа сообщений;

– 3 фаза (с 22 октября по 7 ноября) отличается низким уровнем числа опубликованных записей.

Вторая часть работы посвящена исследованию тональности постов и комментариев для определения уровня агрессии в сообщениях. Разработанная математическая модель [2], основанная на нейронной сети BERT, требует двух этапов обучения. Результатом первого этапа является способность модели работать с корпусом текстов на русском языке. Для второго этапа обучения с учителем необходимо промаркировать часть текстов, которые будет обрабатывать нейросеть, как содержащие или не содержащие некоторый признак. В работе экспертно удалось определить семь типов таких конфликтогенных маркеров. После этого была сформирована обучающая выборка, состоящая из 7 коллекций по 100 сообщений, в которых присутствовал один из семи маркеров, а также из коллекции, содержащей 700 нейтральных сообщений. Оценка качества нейросетевой модели после обучения по этим маркерам показала неудовлетворительный результат и требует дальнейшего исследования. При этом обучение по признаку «конфликтный и неконфликтный текст» позволило выделить собственные маркеры конфликтогенности в тексте и провести разметку сообщений по указанному признаку на удовлетворительном уровне.

Далее был рассчитан процент конфликтных сообщений по отношению к общему числу сообщений. Результаты показали, что за весь период репрезентации конфликта было написано в 1,4 раза больше конфликтных постов по сравнению с неконфликтными. В первую фазу конфликта в постах блогеров преобладает неконфликтная тональность, в постах неинституционализированных СМИ, наоборот, конфликтная. Процент конфликтных комментариев в Telegram-каналах как неинституционализированных СМИ, так и блогеров не зависит от тональности постов и составляет около 60% по отношению ко всем комментариям.

Данная работа показала, что репрезентация арабо-израильского конфликта в российских Telegram-каналах не повлияла на уровень агрессии участников медиасреды. Он оставался постоянным вне зависимости от содержания и тональности постов.

Литература

1. Нефедова Ю. С., Вартанова Е. Л., Аникина М. Е. Математическая модель репрезентации социального конфликта: создание и апробация на корпусе текстов // Медиаскоп. 2024. Вып. 3. URL: <http://www.mediascope.ru/2870>.

2. Нефедов С. И., Вартанов С. А, Рожин А. К. Инженерная реализация математической модели репрезентации социального конфликта // Медиаскоп. 2024. Вып. 3. URL: <http://www.mediascope.ru/2869>.