

УДК 004.89

## МЕТОДОДЫ ОЦЕНКИ ЭМОТИВНОЙ ВАЛЕНТНОСТИ ТЕКСТОВЫХ ВЫСКАЗЫВАНИЙ НА РАЗЛИЧНЫХ ЯЗЫКАХ

Матвеева А.А.

(Университет ИТМО)

Научный руководитель — к.т.н. Махныткина О.В.

(Университет ИТМО)

**Введение.** В данной работе рассмотрены методы оценки эмотивной валентности высказываний дикторов на основе текстовых данных («EmoBERTa», «DialogXL», «DialogueTRM») и проведен их сравнительный анализ.

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР № 622281 «Разработка методов и алгоритмов для мультимодального распознавания валентности высказываний и доминантности дикторов в полилогах».

**Основная часть.** Социальное взаимодействие любого человека невозможно представить без проявления эмоций. Они являются психологической реакцией человека на все что его окружает, выражают предпочтения или неприязнь человека к чему-либо или кому-либо. Эмоции так же могут проявляться и в письменной речи человека. Неудивительно, что, в связи с этим многие компании заинтересованы в распознавании эмотивной валентности текстовых данных для возможности отслеживания предпочтений клиентов их лояльности, оценки качества своего товара и предоставляемого сервиса.

Для проведения сравнительного анализа были выбраны наиболее популярные современные решения по оценке эмотивной валентности, такие как:

- EmoBERTa [1];
- DialogXL [2];
- DialogueTRM [3].

Анализ данных методов оценки эмотивной валентности текстовых высказываний проводился на нескольких наборах данных: «MELD», «IEMOCAP», «MELD RU». «MELD» - набор данных на английском языке. Данный набор данных содержит более 1400 диалогов и 13000 высказываний из сериала «Друзья». Каждое высказывание в диалоге было аннотировано одной из семи эмоций — гнев, отвращение, печаль, радость, нейтральность, удивление и страх.

«MELD RU» - переведенный в автоматическом режиме с помощью MarianMT на русский язык набор данных MELD. «IEMOCAP» - набор данных на английском языке состоит из 151 диалога, каждая реплика в котором аннотируется одной из 9 эмоций (злость, взволнованность, страх, грусть, удивление, фрустрация, счастье, разочарование и нейтральное состояние).

**Выводы.** В данной работе был проведен сравнительный анализ различных методов оценки эмотивной валентности высказываний. После чего с помощью показавшего себя наиболее эффективным среди рассматриваемых аналогов метода были проведены эксперименты по оценке эмотивной валентности высказываний дикторов на основе текстовых данных как на английском, так и на русском языке.

### Список использованных источников:

1. Kim T., Vossen P. EmoBERTa: Speaker-Aware Emotion Recognition in Conversation with RoBERTa //arXiv preprint arXiv:2108.12009. – 2021.
2. Shen W. et al. DialogXL: All-in-one XLNet for multi-party conversation emotion recognition //arXiv preprint arXiv:2012.08695. – 2020.

3. Mao Y. et al. DialogueTRM: Exploring the intra-and inter-modal emotional behaviors in the conversation //arXiv preprint arXiv:2010.07637. – 2020.