

УДК 004.89

ИССЛЕДОВАНИЕ МЕТОДОВ АНАЛИЗА ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА В ДИАЛОГАХ

Мусохранов А.И. (Университет ИТМО),

Научный руководитель – к.т.н., доцент, Махныткина О.В.

(Университет ИТМО)

Введение. Популярность диалоговых агентов растет по мере развития технологий искусственного интеллекта и обработки естественного языка. Важная роль в улучшении работы диалоговых агентов отведена оценке распознавания эмоций человека в разговоре [1]. Исследования выполнены за счет финансирования университета ИТМО в рамках НИР №623088 «Разработка русскоязычного персонифицированного эмоционального диалогового агента».

Основная часть. Задача распознавания эмоций в разговоре направлена на оценку каждого высказывания участника диалога в соответствии с заранее определенным набором эмоций. Первоначально определение эмоционального состояния человека, участвующего в диалоге, происходило в основном с использованием классических методов, таких как выделение ключевых слов, правил, использовании методов машинного обучения. Не учитывался контекст высказываний, предыдущие реплики говорящего.

С развитием технологий искусственного интеллекта глубокие нейронные сети стали широко использоваться для извлечения признаков в задачах распознавания эмоций. Авторы исследования [2] использовали LSTM для извлечения контекстуальных признаков из дикторской речи в задаче анализа настроений. Обученная LSTM не могла как эффективно обобщать новые наборы данных, так и взвешенно оценивать влияние контекста каждого высказывания. В статье [3] предложен модернизированный вариант рекуррентных сетей с механизмом внимания, решавший ранее описанные проблемы. Сеть разговорной памяти (CMN) использовалась для распознавания эмоций в разговоре двух собеседников (диадические диалоги). В CMN отдельные высказывания задействованы в моделировании общего контекста для каждого говорящего. Позднее появилась сеть интерактивной диалоговой памяти (ICON) [4], способная связать высказывания участников разговора для моделирования эмоционального воздействия собеседников друг на друга.

В последнее время набирает популярность Multi-Task Learning (MTL) или обучение с учителем с множественными задачами. MTL представляет собой метод машинного обучения, при котором модель обучается выполнять несколько задач одновременно. Индуктивный перенос предполагает, что знание, накопленное при решении одной задачи, может быть применено для улучшения обучения на другой задаче, даже если задачи не идентичны. Многозадачное обучение смежным задачам может значительно снизить риск переобучения, способствуя повышению обобщения и улучшению качества решения всех подзадач [5, 6]. MTL может быть использовано для комбинирования задач распознавания эмоций с другими задачами, такими как определение сентимента, интенсивности эмоций. В исследовании [7] рассмотрено применение метода Multi-Task Learning для задач распознавания человеческих эмоций и обнаружения сдвига эмоций, т.е. изменения эмоционального состояния собеседника со временем.

Выводы. В результате проведенного анализа исследований на тему распознавания эмоционального состояния человека в диалогах можно сделать вывод, что современные модели классификации эмоций учитывают контекст высказываний, предыдущие реплики собеседников, обладают способностью учитывать влияние собеседников друг на друга.

Список использованных источников:

1. Huang M., Zhu X., Gao J. Challenges in building intelligent open-domain dialog systems // arXiv preprint arXiv:1905.05709, 2020. URL: <https://arxiv.org/pdf/1905.05709.pdf> (дата

обращения: 30.01.2024).

2. Poria S. et al. Context-dependent sentiment analysis in user-generated videos // Proceedings of the 55th annual meeting of the association for computational linguistics, 2017, Vol. 1, pp. 873–883.

3. Hazarika, D., Poria, S., Zadeh, A., Cambria, E., Morency, L. P., & Zimmermann, R. Conversational memory network for emotion recognition in dyadic dialogue videos. In Proceedings of the conference. // Association for Computational Linguistics. North American Chapter. Meeting, 2018, Vol. 1, pp. 2122–2132.

4. Hazarika, D., Poria, S., Mihalcea, R., Cambria, E., & Zimmermann, R. Icon: Interactive conversational memory network for multimodal emotion detection // In Proceedings of the 2018 conference on empirical methods in natural language processing, 2018, pp. 2594-2604.

5. Collobert, R., & Weston, J. A unified architecture for natural language processing: Deep neural networks with multitask learning // In Proceedings of the 25th international conference on Machine learning, 2008, pp. 160-167.

6. McCann B. et al. The natural language decathlon: Multitask learning as question answering //arXiv preprint arXiv:1806.08730, 2018. URL: <https://arxiv.org/pdf/1806.08730.pdf> (дата обращения: 30.01.2024).

7. Gao Q. et al. Emotion recognition in conversations with emotion shift detection based on multi-task learning // Knowledge-Based Systems, Vol. 248, 2022. URL: <https://doi.org/10.1016/j.knosys.2022.108861> (дата обращения: 30.01.2024).