

УДК 004.8

**ДООБУЧЕНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ  
МЕТОДА LOW-RANK ADAPTATION ДЛЯ ГЕНЕРАЦИИ ЭМОЦИОНАЛЬНЫХ  
ОТВЕТОВ ДИАЛОГОВОГО АГЕНТА**

**Павлюкевич В.И. (Университет ИТМО)**

**Научный руководитель – кандидат технических наук Махныткина О.В.  
(Университет ИТМО)**

**Введение.** Исследования выполнены за счет финансирования университета ИТМО в рамках НИР № 623088 «Разработка русскоязычного персонифицированного эмоционального диалогового агента». Среди подходов к дообучению больших языковых моделей, метод Low-Rank Adaptation (LoRA) выделяется своей способностью эффективно адаптировать эти модели к конкретным задачам с минимальными затратами вычислительных ресурсов [1]. Важность этого подхода усиливается в контексте развития диалоговых систем, способных выражать эмоции, что значительно улучшает естественность коммуникации. Представленная работа фокусируется на исследовании применения метода LoRA для дообучения языковых моделей с целью генерации ответов диалогового агента, учитывающих заданную эмоцию.

**Основная часть.** Данная работа направлена на исследование возможности применения метода LoRA для дообучения языковых моделей с целью генерации эмоционально насыщенных ответов в диалоговых системах. LoRA представляет собой передовой подход в области машинного обучения, который позволяет адаптировать большие языковые модели к специфическим задачам, не требуя при этом значительного увеличения числа обучаемых параметров. Это достигается за счет модификации небольшого подмножества весов в существующих архитектурах моделей, таким образом улучшая их способность к адаптации под заданные условия и контексты. В контексте данного исследования LoRA применяется для улучшения эмоциональности ответов диалоговых агентов. Это достигается путем тонкой настройки моделей на специализированных наборах данных, содержащих эмоциональные выражения и контексты. Важным аспектом является выбор и интеграция LoRA с современными архитектурами языковых моделей, такими как Mistral 7B и LLaMa 2 [2, 3]. Это позволяет моделям лучше интерпретировать и генерировать эмоциональные ответы в соответствии с заданным контекстом. Для автоматической оценки эффективности обучения используется классификаторы эмоциональности на базе BERT. [4].

**Выводы.** В рамках проведенного исследования был осуществлен анализ и применение метода Low-Rank Adaptation (LoRA) для дообучения языковых моделей с целью усиления эмоциональности в ответах диалоговых агентов. Применение LoRA показало значительные улучшения в адаптации моделей к специфическим задачам с минимальным увеличением числа обучаемых параметров, что обеспечило не только вычислительную эффективность, но и повышение эмоциональности ответов диалоговых агентов.

**Список использованных источников:**

1. Hu, E., Shen, Y., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W. LoRA: Low-Rank Adaptation of Large Language Models. 2021. <https://arxiv.org/abs/2106.09685>
2. Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., et al. Mistral 7B. 2023. <https://arxiv.org/abs/2310.06825>
3. Touvron, H., Martin, L., Stone, K., et al. Llama 2: Open Foundation and Fine-Tuned Chat Models. 2023. <https://arxiv.org/abs/2307.09288>
4. Zmitrovich, D., Abramov, A., Kalmykov, A., Tikhonova, M., Taktasheva, E., Astafurov, D., Baushenko, M., Snegirev, A., Shavrina, T., Markov, S., Mikhailov, V., Fenogenova, A. A Family of Pretrained Transformer Language Models for Russian. <https://arxiv.org/abs/2309.10931>

Павлюкевич В.И. (автор)

---

(Подпись)

Махныткина О.В. (научный руководитель)

---

(Подпись)