

УДК 004.89

## УЛУЧШЕНИЕ ГЕНЕРАТИВНЫХ СИСТЕМ ОТВЕТОВ НА ВОПРОСЫ СООБЩЕСТВА В ОБЛАСТИ ПРОГРАММИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ ОБРАТНОЙ СВЯЗИ ЛЮДЕЙ

Горбатовский А.В. (Университет ИТМО)

Научный руководитель – кандидат технических наук, Ковальчук С. В.  
(Университет ИТМО)

**Введение.** Основными проблемами генеративных систем, основанных на нейро-сетевом подходе, в области ответов на вопросы сообщества, в частности, в домене программирования, являются ограниченность обобщающей способности текущих систем, а также некорректность и невнимательность систем к деталям вопроса и контекста [1]. В настоящее время нейронные сети, используемые в таких системах, не учитывают человеческую оценку в отношении получаемого ответа. Однако, исследования в отечественном и зарубежном опыте показывают явное улучшение качества и смысловой нагрузки генеративных систем при использовании человеческой обратной связи в процессе обучения [2]. Было установлено, что генеративный подход возможно использовать для проблемы ответа на вопрос. Таким образом, цель данной работы заключается в улучшении генеративных систем ответов на вопросы сообщества в области программирования, путем использования человеческой обратной связи в процессе обучения, с целью достижения более высокой точности и обобщающей способности.

**Основная часть.** Предлагаемое решение состоит из четырёх этапов:

На первом этапе сбора, фильтрации и анализа данных с ресурса Stack Overflow отбираются данные, относящиеся к тегу "python". Исключаются пары вопрос-ответ, содержащие картинки, ссылки и блоки кода.

На втором строится генеративная нейронная сеть для задачи абстрактного ответа на вопрос. Была использована GPT neo, обученная на наборе данных, отобранном и проанализированном на первом этапе. Этот этап позволяет получить начальную модель, способную генерировать ответы на вопросы, которые по человеческой оценке и метрикам качества, таким как BLEU, ROUGE и BertScore, будут лучше, чем первоначальная предобученная сеть. Также данная модель позволит дополнить набор данных для следующего этапа синтетическими данными, полезными для обучения, так как модель имеет изъяны и её ответы можно выделить как негативные примеры для дальнейшего обучения.

На втором этапе строится генеративная нейронная сеть для абстрактного ответа на вопросы, используя модель GPT neo, обученную на наборе данных, отобранном и проанализированном на первом этапе. Полученная модель позволяет генерировать более качественные ответы на вопросы, оцениваемые по метрикам BLEU, ROUGE и BertScore. Также с помощью данной модели можно сгенерировать синтетические данные для дальнейшего обучения, используя её недостатки для выделения негативных примеров.

На третьем этапе строится модель награды, основываясь на улучшенной сети второго этапа. Модель награды представляет регрессионную модель, которая имитирует человеческую оценку сообщества Stack Overflow. Такая модель позволит оценить качество ответа на вопрос, выдаваемое нейронной сетью [3]. Для этого используются различные метрики регрессии (среднее квадратичное отклонение, среднее абсолютное отклонение).

На заключительном этапе используются методы обучения с подкреплением, в частности алгоритм PPO [4], чтобы максимизировать ожидаемую награду от модели генерации. Для того чтобы модель не переобучилась под тренировочный набор данных, она штрафует KL дивергенцией, не сильно отклоняясь от предобученной модели с первого этапа. Возможно также использовать модель награды напрямую, в таком случае ожидаемая награда генеративной сети может не являться достоверной оценкой качества ответа, что свидетельствует о переобучении.

Комбинируя все три этапа, получаем предложенное решение, которое основывается на

использовании генеративной нейронной сети GPT neo, модели награды и методах обучения с подкреплением. Такой подход позволяет улучшить обобщающую способность нейронной сети, учитывать человеческую оценку качества ответов и повысить точность генерации ответов на вопросы в области программирования.

**Выводы.** Разработаны и протестированы методы для улучшения генеративных систем ответов на вопросы в области программирования. Было показано, что использование обратной связи от сообщества может значительно улучшить качество ответов, а методы обучения с подкреплением позволяют достигать более высокой точности, последовательности и осмысленности ответов так же, как и улучшение целевых метрик генерации текста. Полученные результаты могут быть применены для создания более эффективных систем ответов на вопросы в других областях.

#### **Список использованных источников:**

1. Kovalchuk S. V., Lomshakov V., Aliev A. Human perceiving behavior modeling in evaluation of code generation models //Proceedings of the 2nd Workshop on Natural Language Generation, Evaluation, and Metrics (GEM). – 2022. – P. 287-294.
2. Ouyang L. et al. Training language models to follow instructions with human feedback //arXiv preprint arXiv:2203.02155. – 2022.
3. Shah C., Pomerantz J. Evaluating and predicting answer quality in community QA //Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. – 2010. – P. 411-418.
4. Schulman J. et al. Proximal policy optimization algorithms //arXiv preprint arXiv:1707.06347. – 2017.