

TELEGRAM-БОТ ДЛЯ ИЗУЧЕНИЯ АНГЛИЙСКОГО: ГЕНЕРАЦИЯ УПРАЖНЕНИЙ С ИСПОЛЬЗОВАНИЕМ ЯЗЫКОВОЙ МОДЕЛИ DEEPSEEK

А. Д. Лезина, В. И. Мелкозерова, А. В. Фаттахова

НИУ «Высшая школа экономики»

Санкт-Петербург

В настоящем исследовании предлагается новый подход к реализации Telegram-бота для помощи в изучении английского языка через интерактивные задания. Бот учитывает, какой у пользователя уровень владения языком и позволяет практиковать три ключевых аспекта языка: перевод, лексику и грамматику. Для упражнений всех уровней доступны разные лексические и грамматические темы. Для генерации упражнений используется большая языковая модель DeepSeek-R1, производящая задания в реальном времени, что способствует повышению динамичности обучения. Промпт-тюнинг позволяет гибко настроить модель для генерации задач без необходимости переобучения, снизить затраты на вычислительные ресурсы, а также быстро адаптировать модель под новые темы или типы упражнений, изменяя только промпты. Основные преимущества реализованного ассистента заключаются в персонализации (выбор уровня сложности и темы), привлечении синтетических данных (задания генерируются автоматически, что экономит время и ресурсы), а также в интерактивности (формат чата делает процесс обучения удобным и увлекательным). Среди ограничений можно выделить галлюцинации, возникающие при генерации текста на русском языке и ограниченность типов упражнений. Наш проект доступен с открытым исходным кодом, что предоставляет возможность для его модификации и совершенствования: <https://github.com/AnnaFattakhova/AnnaFattakhova/tree/main/Telegram-bot>.

Ключевые слова: обучение иностранному языку, Telegram-бот, большая языковая модель, генерация упражнений

TELEGRAM BOT FOR LEARNING ENGLISH: GENERATING EXERCISES USING THE DEEPSEEK LANGUAGE MODEL

A. D. Lezina, V. I. Melkozerova, A. V. Fattakhova

National Research University Higher School of Economics

St. Petersburg

This study proposes a new approach to implementing a Telegram bot to assist in learning English through interactive exercises. The bot considers the user's language proficiency level and allows practicing translation, vocabulary, and grammar. For exercises at all levels, various lexical and grammatical topics are available. To generate exercises, the large language model DeepSeek-R1 is used, producing tasks in real time, which enhances the dynamism of learning. Prompt tuning enables flexible model adaptation for generating tasks without retraining, reduces computational resource costs, and allows quick adjustments by modifying the prompts. The advantages of the implemented assistant include personalization, the use of synthetic data, and interactivity. Among the limitations there are hallucinations that occur when generating text in Russian and the limited variety of exercise types. Our project is open-source: <https://github.com/AnnaFattakhova/AnnaFattakhova/tree/main/Telegram-bot>.

Keywords: second language learning, Telegram-bot, large language model, exercise generation

Развитие технологий генеративного искусственного интеллекта привело к активному внедрению больших языковых моделей в сферу образования. Наиболее эффективные помощники для изучения иностранных языков сейчас – чат-боты в роли ассистентов. Несмотря на большое количество существующих решений, многие боты не адаптированы к разным уровням владения языком и не предоставляют пользователю качественную обратную связь.

Цель исследования – оценить полезность и удобство разработанного нами бота и выявить направления потенциальной доработки. В ходе анализа выявлены ключевые аспекты, требующие улучшения, среди которых разнообразие заданий и персонализация, а также предложения по оптимизации функционала.

По определению Г. Д. Хуанга и Ч. Ю. Чанга, чат-бот – модель, поддерживающая межличностное общение, облегчающая обучение и предоставляющая информацию через интерактивные и удобные интерфейсы [1]. Использование чат-ботов в обучении языкам становится популярным, поскольку большие языковые модели способны вести диалог, приближенный к диалогу на естественном языке. Как отмечают

специалисты, интерес к исследованиям в этой области значительно вырос с 2020 г., а к 2022 г. количество работ растет в геометрической прогрессии [2].

Преимущества чат-ботов – простота использования, оперативность, возможность персонализации и повышение мотивации [3; 4]. Из недостатков стоит выделить ограничения по уровню владения языком (A2–B1) и когнитивную нагрузку (избыток информации) [3; 5].

Несмотря на растущий интерес, массового использования чат-ботов в языковом образовании пока нет, более того, недостаточно эмпирических данных, подтверждающих их пользу для учащихся [3]. Чтобы устранить этот пробел, необходимы исследования эффективности чат-ботов.

В рамках проекта был реализован бот в Telegram, предназначенный для практики основных аспектов английского языка. Принцип его работы: сначала пользователь выбирает категорию упражнений («Перевод», «Лексика» или «Грамматика»), затем уровень владения языком для генерации задания соответствующей сложности. На этапе тестирования доступны начальный (A1–A2), средний (B1–B2) и продвинутый (C1–C2) уровни; это упрощает выбор для пользователей, не знакомых с тонкостями разделения на подуровни. Однако в дальнейшем планируется ввести не только шесть основных категорий, но и смежные (например, B2–C1).

Далее предоставлен выбор темы внутри категории:

- «Лексика» и «Перевод»: Еда, Погода, Путешествия, Хобби, Культура, Здоровье;
- «Грамматика»: Времена, Артикли, Предлоги, Условные предложения, Пассивный залог, Модальные глаголы.

Мы выбрали наиболее популярные лексические и грамматические темы, содержащиеся в каждом учебном пособии по английскому языку. Важно было отобрать широкие темы, подходящие для изучения почти на всех уровнях владения, так как механика выбора темы в зависимости от уровня еще не реализована.

В категории «Перевод» модель генерирует текст по выбранной теме на русском языке, который пользователь должен перевести на английский. Для «Грамматики» создается задание с пропуском и тремя вариантами ответа, один из которых верный. При выборе «Лексики» модель сначала генерирует словарь по выбранной теме из 15 английских слов с их переводом на русский, после чего создает задание с пропуском и тремя вариантами ответа с лексикой из словаря.

У пользователя всегда есть возможность при помощи кнопок вернуться и выбрать другую категорию или уровень. Код, примеры заданий и работы кнопок доступны по ссылке: <https://github.com/AnnaFattakhova/AnnaFattakhova/tree/main/Telegram-bot>. На рис. 1 логика работы бота представлена в виде блок-схемы.

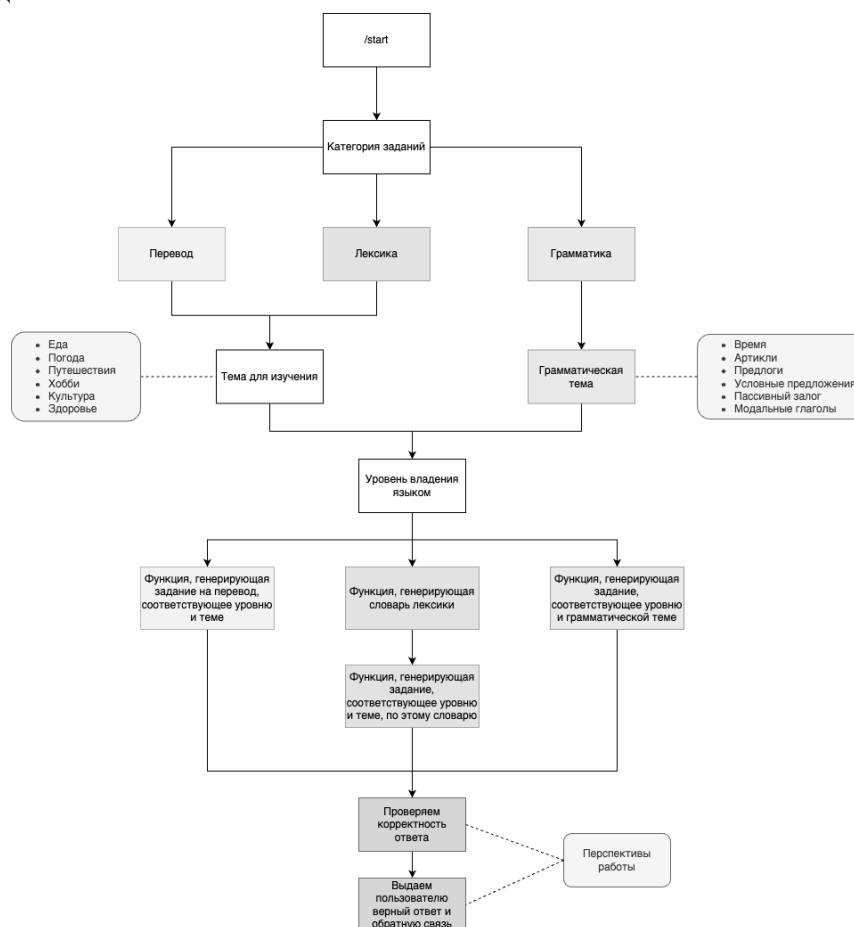


Рис. 1. Блок-схема работы алгоритма

Для генерации заданий в режиме реального времени с помощью промт-тюнинга используется языковая модель DeepSeek-R1-Distill-Qwen-32B, взаимодействие с которой реализовано через API. Выбор модели обусловлен ее доступностью для пользователей, так как мы разрабатываем продукт для обучающихся в Российской Федерации. Модель поддерживает как русский, так и английский язык, а ее контекстное окно (128 тыс. токенов) достаточное, чтобы генерировать задания любого типа. В перспективе мы планируем провести тесты с разными моделями, чтобы сравнить их производительность и выбрать лучший вариант для нашего бота.

Для оценки разработанного продукта была собрана обратная связь от потенциальных пользователей с помощью анкеты в Google Forms. В исследовании приняло участие 35 респондентов, разделенных на три возрастные группы: 18–21 («студенты») – 15 чел., 22–28 (молодежь) – 16 чел. и 52–64 (третий возраст) – 4 чел. Уровень владения английским среди опрошенных: 42% владеют языком на среднем уровне (B1–B2), 34,3% определяют свой уровень как продвинутый (C1–C2) и 22,9% владеют основами английского (A1–A2).

Респондентам было представлено описание функционала Telegram-бота, а также продемонстрированы скриншоты с примерами его работы. Участникам предлагалось оценить пять утверждений по шкале от 1 (однозначно нет) до 5 (однозначно да). Ниже представлены вопросы, а также распределение оценок респондентов по шкале (рис. 2):

1. Этот бот полезен для изучения английского языка.
2. Этот бот содержит достаточно разнообразных тем для изучения английского языка на разных уровнях.
3. Этот бот содержит достаточно разнообразных типов заданий для изучения английского языка на разных уровнях.
4. У этого бота достаточно понятный и удобный в использовании интерфейс.
5. Какова вероятность того, что вы стали бы использовать такой бот для изучения или преподавания английского языка?

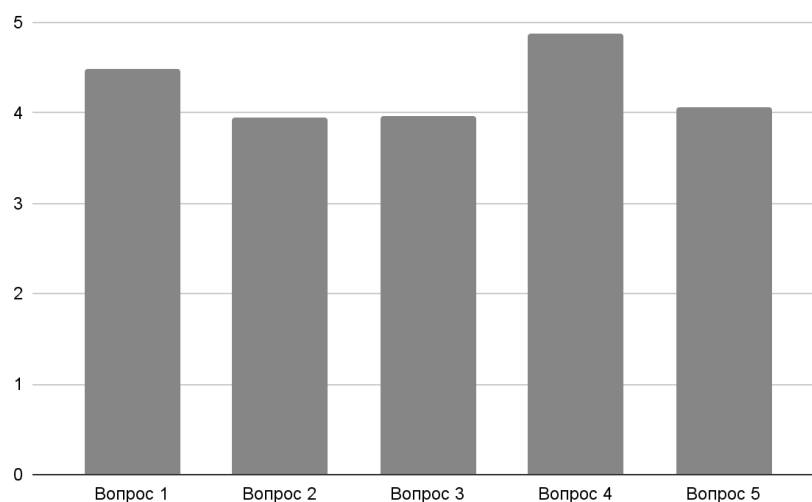


Рис. 2. Распределение оценок респондентов по каждому из вопросов

Мы проанализировали связь между принадлежностью респондентов к различным категориям и их ответами (табл.). Наиболее высокие оценки по всем вопросам выставили участники из третьей возрастной группы (52–64). Респонденты первой группы (18–21 год) высоко оценили удобство интерфейса, тогда как представители второй группы (22–28 лет) отметили полезность, разнообразие тем и типов заданий.

Таблица. Среднее (μ) и среднеквадратическое отклонение (σ) по каждому из вопросов

Вопрос		18–21	22–28	52–64	A1–A2	B1–B2	C1–C2	Учат язык	Преподают язык
№ 1 (Полезность)	μ	4,27	4,56	5	4,88	4,4	4,3	4,53	4,4
	σ	0,8	0,73	0	0,35	0,74	0,89	0,74	0,89
№ 2 (Разнообразие тем)	μ	3,73	3,88	5	4,5	3,8	3,75	4,07	3,4
	σ	1,03	1,15	0	0,76	1,37	0,75	0,96	0,89
№ 3 (Разнообразие типов заданий)	μ	3,73	4	4,75	4,63	3,8	3,75	3,93	3,8
	σ	1,22	1,15	0,5	0,52	1,42	0,97	1,03	1,1
№ 4 (Удобство интерфейса)	μ	4,93	4,81	5	5	4,93	4,75	4,93	5
	σ	0,26	0,40	0	0	0,26	0,45	0,26	0
№ 5 (Вероятность использования)	μ	3,8	4,06	5	4,75	4,13	3,5	4,07	4,2
	σ	1,42	1,18	0	0,46	1,06	1,62	1,33	1,1

Выявлена обратная зависимость между средним количеством баллов, выставяемых респондентами по каждому из вопросов, и их уровнем владения языком (рис. 3). Из графика следует, чем выше уровень владения английским, тем менее удобным и полезным респондентам представляется продукт. Это говорит о необходимости внести разнообразие в типы заданий и добавить более сложные лексические темы.

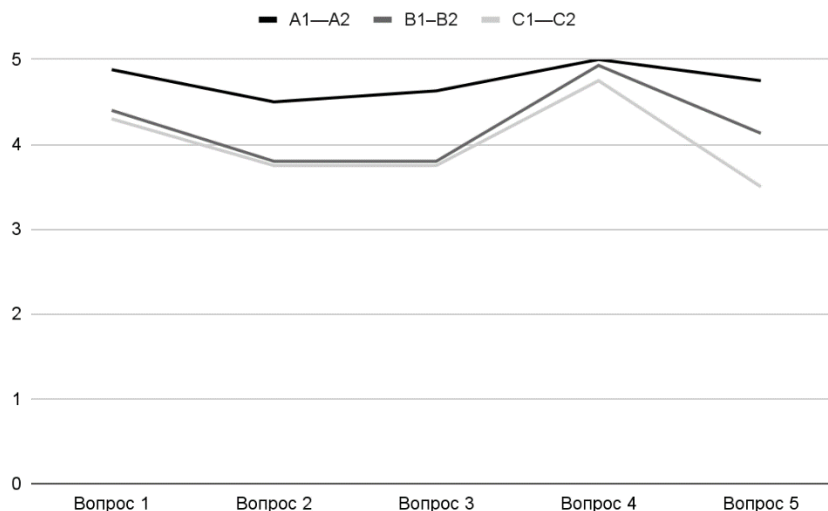


Рис. 3. Распределение оценок респондентов по каждому из вопросов в зависимости от уровня владения языком

В дополнение к количественным данным были собраны качественные отзывы участников опроса. Респонденты отмечали хорошую идею и удобный интерфейс, а также выдвинули несколько интересных предложений по доработке бота:

- возможность получить пояснение или правила для заданий;
- темы «Одежда», «Современные технологии», «Экология»;
- аудио- и видеозадания;
- рубрики «Слово дня» и «Повторение пройденного».

Комплексный анализ количественных и качественных показателей продемонстрировал заинтересованность потенциальных пользователей, но также указал на необходимость доработки, особенно в разнообразии тем и типов заданий. После оптимизации упомянутых аспектов планируем провести дополнительные тестирования, чтобы выяснить, насколько использование нашего продукта помогает пользователям углубить понимание конкретных лексических и грамматических тем.

На данном этапе можно выделить несколько ограничений проекта. Во-первых, качество генерации синтетического материала для заданий напрямую зависит от возможностей модели DeepSeek: иногда при генерации текста на русском языке возникают галлюцинации. Для минимизации проблемы необходимо разработать более сложные механизмы контроля: подключение правил фильтрации, внедрение семантического поиска и валидации выходных данных, что позволит не включать в вывод Reasoning модели.

Кроме того, на этапе разработки находится функция, которая позволит пользователям получать развернутую обратную связь от бота. Она будет состоять из проверки корректности (анализ ответа, предоставление решения с пояснениями) и мотивации (сообщение с похвалой или поддержкой, советами). Такой алгоритм позволит повысить персонализированность и интерактивность, что сделает бота более эффективным для образовательных целей.

Итак, бот выполняет поставленные задачи по генерации упражнений для изучения английского языка по запросу пользователя, но в настоящий момент есть ряд ограничений, затрудняющих его эффективную работу. Добавление таких функций, как проверка ответов с обратной связью и отслеживанием прогресса, генерация справочной информации, а также расширенный список тем и типов заданий, сделает бота более функциональным и персонализированным инструментом для изучения английского.

Главные преимущества бота – экономия ресурсов за счет автоматической генерации задач и гибкость настройки без переобучения. Практическая ценность заключается в мобильности, так как бот доступен в Telegram с любого устройства, и возможности подстроиться под уровень знаний и интересы пользователя. Таким образом, бот представляет собой эффективный и доступный инструмент для изучения английского языка в эпоху цифровых технологий.

ЛИТЕРАТУРА

1. Hwang G. J., Chang C. Y. A Review of Opportunities and Challenges of Chatbots in Education // *Interactive Learning Environments*, 2023. Vol. 31, № 7. P. 4099–4112.
2. Jeon J., Lee S., Choi S. A Systematic Review of Research on Speech-Recognition Chatbots for Language Learning: Implications for Future Directions in the Era of Large Language Models // *Interactive Learning Environments*, 2024. Vol. 32, № 8. P. 4613–4631.
3. Huang W., Hew K. F., Fryer L. K. Chatbots for Language Learning – Are They Really Useful? A Systematic Review of Chatbot-Supported Language Learning // *Journal of Computer Assisted Learning*, 2022. Vol. 38, № 1. P. 237–257.
4. Прохорова А. А., Катаева А. Л. Использование чат-ботов при обучении иностранному языку учащихся средней школы // *Проблемы романо-германской филологии, педагогики и методики преподавания иностранных языков*, 2024. № 20. С. 214–221.
5. Сысоев П. В., Филатов Е. М. Чат-боты в обучении иностранному языку: преимущества и спорные вопросы // *Вестник ТГУ*, 2023. № 1. С. 66–72.