

СЕРВИС ДЛЯ ПРЕПОДАВАТЕЛЕЙ И ОБУЧАЮЩИХСЯ ПО ТЕМАТИЧЕСКОМУ ТЕСТИРОВАНИЮ ЗНАНИЙ АНГЛИЙСКОГО ЯЗЫКА ОСНОВАННЫЙ НА ГЕНЕРАТИВНЫХ МОДЕЛЯХ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА

Лайок О. В. (Университет ИТМО)

Научный руководитель – доцент факультета ИКТ, кандидат технических наук

Добренко Н. В.

(Университет ИТМО)

Введение. В настоящее время в сфере образования существует тренд на персонификацию учебного процесса, и обучение английскому языку не является исключением. Для разработки индивидуального плана ученика за последние годы было составлено множество методических материалов и пособий, а также набирают популярность сервисы, которые берут на себя часть задач учителя и позволяют ученику самому изучать выбранную область. И хотя предложенные варианты во многом решают проблему качественного персонифицированного образования, некоторые разделы обучения английскому языку так и остаются без внимания, в частности создание индивидуальных тестов по пройденным темам для проверки усвоенных знаний. В моей работе рассматривается вариант сервиса, который позволит решить данную проблему.

Основная часть. Для написания уникальных тестовых материалов предлагается использовать генеративную модель обработки естественного языка. Такой тип модели предобучен на огромном корпусе документов, что позволяет создать уникальный текст по заданной теме [1]. В последние годы в сфере был произведен прорыв в технологиях (например, исследования компаний OpenAI [2] и Google [3]), который расширил круг задач, решаемых алгоритмами обработки естественного языка. Для поставленной мной проблемы прекрасно подходит данная модель из-за следующих преимуществ:

1. Модель может понять запрос, написанный на естественном языке и выдать осмысленный результат, в частности создать тест в заданной форме по заданной теме.
2. Уникальность выдаваемого результата близка к идеальной, что помогает решить проблему списывания или плагиата других пособий.
3. Использование результатов работы модели преподавателем в разы снижает затраченное время на подготовку к занятиям, что положительно сказывается на качестве его работы.
4. Модель, настроенная на решение данной задачи способна создавать тестовые материалы не только по различным темам, но и в разных форматах, что позволяет не только использовать один сервис для решения множества задач, но и создать более полную картину для индивидуализации обучения пользователя.

В ходе исследования существующих архитектур и вариантов предобученных моделей данного класса был произведен сравнительный анализ существующих решений, на основе которых можно создать требуемую интеллектуальную систему. В результате анализа по критериям доступности, точности результатов, легкости в имплементации и дообучении модели было принято решение остановиться на модели GPT-3. Был разработан прототип сервиса на основе чат-бота, который позволяет по заданной теме и форме теста получить готовый материал для проверки знаний по интересующей пользователя теме. В качестве места размещения данной системе был выбран вариант чат-бота в мессенджере Telegram.

Выводы. Материалы, созданные выбранным алгоритмом, были протестированы преподавателями и учениками, которые положительно отозвались об их применении в образовательном процессе, что подтверждает гипотезу о целесообразности создания такой информационной системы. При этом был сформулирован ряд правок и предложений по

улучшению качества генерируемых материалов, что может быть использовано в ходе развертки данного сервиса и подготовке к публичному использованию.

Список использованных источников:

1. Kalyan K. S. AMMUS : A Survey of Transformer-based Pretrained Models in Natural Language Processing / Rajasekharan A., Sangeetha S. [Электронный ресурс] // arxiv.org. – 2021. – URL: <https://arxiv.org/abs/2108.05542> (дата обращения 18.02.2023)
2. Training language models to follow instructions with human feedback / Ouyang L., Jeff Wu, Xu Jiang [и др.] [Электронный ресурс] // arxiv.org. – 2022 – URL: <https://arxiv.org/abs/2203.02155> (дата обращения 18.02.2023)
3. LaMDA: Language Models for Dialog Applications / Thoppilan R., Daniel De Freitas, Hall J. [и др.] [Электронный ресурс] // arxiv.org. – 2022 – URL: <https://arxiv.org/abs/2201.08239> (дата обращения 18.02.2023)