

УДК 004.8

ПРИМЕНЕНИЕ МОДЕЛЕЙ СЕМЕЙСТВА BERT В ЗАДАЧАХ HR-КОММУНИКАЦИИ

Д. В. БАДРЕТДИНОВ

Научный руководитель Е. А. ЗАЙЧЕНКО

Белорусско-Российский университет

Могилев, Беларусь

Ключевым элементом цифровизации HR является использование технологий искусственного интеллекта (ИИ) в поиске и подборе персонала. Применение ИИ повышает эффективность, улучшает опыт кандидатов и позволяет масштабировать процессы найма без увеличения нагрузки на сотрудников. Данная работа посвящена разработке и исследованию практического применения моделей семейства BERT для решения задач классификации сообщений соискателей работы. Проведен детальный анализ различных архитектур моделей BERT, таких как google/rembert, FacebookAI/xlm-roberta-large, sentence-transformers/LaBSE и ai-forever/ruRoberta-large, что позволило выбрать наиболее подходящую модель для русскоязычной коммуникации – ai-forever/ruRoberta-large.

Работа включала проведение серии экспериментов по оптимизации объёма обучающих данных, которые показали, что достаточно 150–300 примеров на класс для достижения оптимальной точности без переобучения модели. Выявлена зависимость качества классификации от количества обучающих примеров и определена точка насыщения, после которой прирост точности становится незначительным.

Кроме того, выполнен сравнительный анализ различных подходов параллельного запуска моделей BERT на GPU (синхронного, асинхронного и многопоточного), который подтвердил, что синхронный подход является наиболее эффективным с точки зрения производительности и потребления ресурсов.

Также было проведено исследование экономической эффективности использования различных GPU (RTX 6000 Ada, L40S, H100 NVL и др.). На основе этого анализа для практической реализации задачи классификации была выбрана видеокарта RTX 6000 Ada, обеспечивающая оптимальное соотношение производительности и стоимости.

Практическое применение разработанной модели реализовано в виде двухуровневой архитектуры: первый уровень осуществляет общую классификацию сообщений по их типу. В результате проведенных экспериментов достигнута высокая точность классификации – до 100 % на первом уровне и 92 %...93 % на втором.

Таким образом, предложенное решение демонстрирует высокую эффективность и применимость для автоматизации обработки сообщений соискателей в HR-коммуникации, обеспечивая значительное ускорение и повышение качества работы с кандидатами.