Греков Максим Александрович, магистрант,
Российский новый университет
Grekov Maxim Aleksandrovich,
Russian New University

СИСТЕМА АВТОМАТИЗИРОВАННОГО ОБУЧЕНИЯ ЯЗЫКОВОЙ МОДЕЛИ НА ТЕКСТАХ ТРУДОВОГО КОДЕКСА РОССИЙСКОЙ ФЕДЕРАЦИИ

Аннотация. В статье представлена разработанная система для дообучения крупной языковой модели LLaMA-3 на тексте Трудового кодекса Российской

Федерации. Система оптимизирована для работы на оборудовании с ограниченными вычислительными ресурсами. Описаны методы предобработки юридических текстов, параметры обучения с использованием технологии QLoRA, а также проведена оценка качества полученной модели. Разработанное решение позволяет создавать специализированных юридических ассистентов для автоматизации работы с трудовым законодательством.

Ключевые слова: Искусственный интеллект, языковые модели, тонкая настройка, Трудовой кодекс, QLoRA, юридические технологии.

Введение

Современные большие языковые модели демонстрируют выдающиеся способности в обработке естественного языка, однако их эффективность в специализированных доменах, таких как юриспруденция, требует дополнительной адаптации [1]. Трудовой кодекс Российской Федерации представляет собой сложный структурированный нормативный акт, содержащий специфическую терминологию и строгие формулировки, что обуславливает необходимость специализированного обучения моделей для работы с данным типом документов [2].

Основной проблемой при дообучении крупных языковых моделей является требование к значительным вычислительным ресурсам. Традиционные методы тонкой настройки требуют наличия специализированного оборудования с большим объемом видеопамяти, что ограничивает доступность технологий для исследовательских организаций и образовательных учреждений [3].

В данной работе представлена система, позволяющая эффективно дообучать модель LLaMA-3 на текстах Трудового кодекса РФ с использованием потребительского оборудования. Система реализует поэтапный подход к обучению с применением методов параметрически-эффективной тонкой настройки QLoRA [4].

Методология

Для решения поставленной задачи была разработана модульная система, включающая следующие компоненты: модуль предобработки данных, систему поэтапного обучения и механизм объединения результатов.

Предобработка данных осуществлялась с использованием специализированного парсера, разработанного для извлечения структурированной информации из текста Трудового кодекса. В процессе обработки выделялись статьи, главы и разделы кодекса, которые преобразовывались в набор данных формата "инструкция-ответ". Всего было создано более 400 обучающих примеров, охватывающих основные положения трудового законодательства [5].

Архитектура системы обучения основана на модели LLaMA-3 с 8 миллиардами параметров, использующей 4-битное квантование для уменьшения требований к памяти. Для эффективного обучения на оборудовании с 8 ГБ видеопамяти применялась технология QLoRA

(Quantized Low-Rank Adaptation), позволяющая осуществлять тонкую настройку с минимальными вычислительными затратами [4].

Параметры обучения были оптимизированы для видеокарты RTX 4060: скорость обучения 2e-4, размер пакета 1, ранг матрицы адаптации 8. Система реализует поэтапное обучение, при котором данные разбиваются на батчи, что позволяет обойти ограничения по памяти без потери качества итоговой модели.

Результаты и обсуждение

Разработанная система успешно прошла полный цикл обучения, продемонстрировав стабильную работу в условиях ограниченных вычислительных ресурсов. Качественная оценка результатов показала значительное улучшение способностей модели в области трудового права по сравнению с исходной версией.

Обученная модель продемонстрировала способность точно цитировать статьи Трудового кодекса РФ и давать развернутые ответы на юридические вопросы. Например, при запросе о понятии трудового договора модель не только давала точное определение, но и ссылалась на соответствующую статью законодательства, что свидетельствует о successful усвоении предметной области.

Сравнительный анализ ответов исходной и обученной моделей выявил следующие преимущества дообученной версии: точность цитирования нормативных положений, полнота предоставляемой информации, соответствие ответов юридическому стилю и минимальное количество галлюцинаций. Эти результаты подтверждают эффективность выбранного подхода к подготовке данных и организации процесса обучения.

Важным достижением является разработанная система поэтапного обучения, которая позволяет обойти ограничения оборудования без существенного ущерба для качества модели. Механизм объединения обученных батчей обеспечивает создание целостной модели, готовой к практическому применению.

Заключение

Представленная в работе система демонстрирует эффективный подход к адаптации крупных языковых моделей для специализированных юридических задач. Использование методов QLoRA и поэтапного обучения позволяет существенно снизить требования к вычислительным ресурсам, делая технологию доступной для широкого круга пользователей.

Полученные результаты подтверждают перспективность разработанного решения для создания специализированных юридических ассистентов. Система может быть интегрирована в образовательные платформы, корпоративные HR-системы и юридические консультационные сервисы, способствуя автоматизации работы с трудовым законодательством.

Дальнейшие исследования могут быть направлены на расширение функциональности системы за счет включения дополнительных нормативных актов и совершенствования методов оценки качества генерации юридических текстов.

Список литературы:

1. Touvron H. et al. LLaMA: Open and Efficient Foundation Language Models
2. Dettmers T. et al. QLoRA: Efficient Finetuning of Quantized LLMs
3. Brown T.B. et al. Language Models are Few-Shot Learners
4. Hu E. J. et al. LoRA: Low-Rank Adaptation of Large Language Models
5. Трудовой кодекс Российской Федерации