

## СЕРВИС АВТОМАТИЧЕСКОЙ МОДЕРАЦИИ РЕЗЮМЕ НА РУССКОМ ЯЗЫКЕ

Шилонос В.Р. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Федоров Д.А. (ИТМО)

**Введение.** Модерация резюме является актуальной задачей для платформ по размещению вакансий и резюме, заказчик, комитет по труду и занятости населения Санкт-Петербурга, сталкивается с потоком из нескольких сотен резюме в день, каждое из которых должно быть тщательно просмотрено и промодерировано. Сейчас эту работу выполняют сотрудники комитета, но это очень трудоемкая и монотонная работа, поэтому комитет выразил желание автоматизировать этот процесс. Автоматизация модерации может быть реализована на основе современных подходов в обработке естественного языка и больших языковых моделей.

**Основная часть.** Комитет предоставил документы с требованиями к резюме и данные резюме с разметкой, приняты ли они к публикации, на этих данных были протестированы следующие подходы

1. Анализ настроения текста резюме - была использована модель RuBert Base Russian Case Russian Sentiment [1] для определения настроения текста, чтобы затем считать не прошедшими модерацию тексты с сильным негативным или позитивным настроением, и считать прошедшими резюме с нейтральным настроением.

2. Сравнение желаемой должности и опыта кандидата - была выбрана модель T-pro-it-1.0 [2], которая переводит текст желаемой должности и опыта в одно векторное пространство, затем в нём находилась метрика - косинусное расстояние между желаемой должностью и опытом работы, которая обозначала близость по смыслу желаемой должности и опыта, чтобы затем считать резюме, где опыт не похож на желаемую должность, не прошедшими модерацию.

3. Модерация предобученной моделью - тексты резюме были даны на вход большой языковой модели Llama Guard 3 8B [3] предобученной на классификацию контента на опасный и безопасный контент.

4. Zero-Shot модерация на основе правил - на основе документов требований к резюме был сформирован формальный список правил, по которым модерируются резюме, затем этот список был передан большой языковой модели T-pro-it-1.0 [2], вместе с текстом резюме, после чего модели было дано указание проверить резюме на соответствии каждому правилу, после чего вернуть все правила, которым противоречит

**Выводы.** Проведено апробирование подходов обработки естественного языка и больших языковых моделей для задачи модерации резюме на данных предоставленных комитетом по труду и занятости населения Санкт-Петербурга.

### Список использованных источников:

1. RuBert Base Russian Case Russian Sentiment // huggingface URL: <https://huggingface.co/r1char9/rubert-base-cased-russian-sentiment> (дата обращения: 13.01.2025).
2. T-pro-it-1.0 // huggingface URL: <https://huggingface.co/t-tech/T-pro-it-1.0> (дата обращения: 13.01.2025).
3. Llama-Guard-3-8B // huggingface URL: <https://huggingface.co/meta-llama/Llama-Guard-3-8B> (дата обращения: 13.01.2025).