

# НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

## «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

### Факультет гуманитарных наук

#### Образовательная программа «Фундаментальная и компьютерная лингвистика»

#### Отзыв рецензента на выпускную квалификационную работу

**Эдуарда Геннадьевича Григорьева, студента 4 курса бакалавриата группы  
БКЛ162 факультета гуманитарных наук,  
на тему: «Обнаружение пропаганды с использованием нейронных сетей в  
новостных текстах на русском языке»**

Научная проблема, которую Эдуард Геннадьевич Григорьев выбрал для исследования в своей выпускной квалификационной работе, лежит на пересечении области лингвистической прагматики и информационного поиска. Цель исследования состоит в том, чтобы на основе четко определенных лингвистических приемов речевого воздействия, характеризующих пропаганду как коммуникативную стратегию, автоматически выявлять в текстах новостных сообщений фрагменты, по которым тексты следует рассматривать как пропагандистские.

В исследовательской работе автор подробнейшим образом рассматривает как явление пропаганды, так и основные понятия и методики, необходимые для его научного анализа. В работе представлен обзор пропагандистских техник, из которых 680казывается в фокусе исследования, для каждой из выбранных техник приводятся иллюстративные примеры. Автор противопоставляет лживые и неточные утверждения пропагандистским высказываниям. В связи с этим у рецензента возник вопрос: **поскольку пропаганда – это одна из техник речевого манипулирования, каким образом она соотносится, например, с рекламой?**

Для проведения экспериментов автор подготовил и разметил набор данных, включающий 54 статьи из 10 источников. Разнообразие источников, авторов и тем текстов обеспечивает репрезентативность корпуса. Соглашусь с тем, что набор данных небольшой, но поставленная задача требует его тщательной подготовки, поэтому автор ВКР прав в том, что не стремился на данном этапе увеличить объем данных, а сосредоточился на его качественном анализе. Набор данных был размечен автором вручную. Хотелось бы отметить следующее обстоятельство: при работе с текстами СМИ надо учитывать тот факт, что многие из авторов новостных сообщений работают в разных жанрах журналистики, кроме того, велика доля новостных сообщений, которые являются транскриптами устных репортажей. В частности, в наборе данных есть тексты интервью (близких к устной коммуникации). Вопрос заключается в том, **есть ли различие между применением пропагандистских техник в текстах письменных или предназначенных для устной коммуникации?**

Количественный анализ данных позволил соотнести пропагандистские техники и основные лингвостатистические параметры текстов, а также выявить распределение различных источников информации по предпочтительному использованию тех или иных техник. В итоге автор принял решение выделить 18 техник, которые можно быстро распознавать автоматическими методами. Для этого автор выбрал

нейросетевую модель BERT в нескольких вариантах, обоснованно выбрал параметры обучения и оценки качества моделей. Рассматривались стандартная модель BERT, объединенная модель BERT, гранулярная модель BERT и мультигранулярная модель BERT с двумя функциями активации, ReLU и Sigmoid. Автор приводит все данные об оценке качества распознавания пропаганды с помощью моделей, в экспериментах удалось достигнуть высокого для такой задачи результата,  $F1=0,64$ . Вопрос рецензента касается того, **есть ли альтернативы BERT на сегодняшний день, или по существу, это единственно возможное решение?**

Исследование, проведенное Эдуардом Геннадьевичем Григорьевым, нашло воплощение в самостоятельном сервисе автоматического распознавания пропаганды, работающего с четырьмя моделями в двух основных режимах (анализ случайного текста и анализ текста, предложенного пользователем). **Планирует ли автор использовать его для расширения исходного набора данных, и если да, каким образом?**

Вопросы, сформулированные в рецензии, носят дискуссионный характер, и возможные ответы автора смогут пролить свет на детали, могущие оказаться ценными с научной точки зрения. Замечания по ВКР отсутствуют.

Итак, выпускная квалификационная работа, представленная к защите, соответствует всем требованиям, предъявляемым к научным произведениям такого уровня, а ее автор заслуживает присвоения квалификации бакалавра лингвистики. Содержание ВКР полностью соответствует заявленной в названии теме, тема раскрыта полностью, структура ВКР соответствует цели и задачам исследования, в ВКР в полной мере освещены актуальные проблемы теоретического и практического характера. При выполнении исследования и подготовке текста ВКР автор использовал современную литературу, достижения науки и практики. Текст ВКР удобным образом структурирован, написан научным стилем. Автор четко формулирует теоретические предпосылки исследования, детально комментирует практический материал, дает развернутое обоснование выводов. Результаты, достигнутые автором ВКР, имеют высокую теоретическую значимость, практическую ценность и могут найти применение в коммуникологии как таковой и в прикладных областях знаний, таких как мониторинг новостных потоков, социальных сетей и разработка диалоговых систем.

Рецензент поздравляет автора и научного руководителя с успешным завершением исследования и желает дальнейших научных и творческих удач!

Рецензент

Митрофанова Ольга Александровна,  
кандидат филологических наук, доцент, доцент,  
кафедра математической лингвистики СПбГУ

Дата

3 июня 2020 года