

## ОТЗЫВ

на автореферат диссертации Смирнова Ивана Валентиновича «Методы разноуровневого анализа текстов на естественном языке и их приложения в системах информационного поиска и психолингвистических исследованиях», представленной на соискание ученой степени доктора технических наук по специальности 1.2.1 – «Искусственный интеллект и машинное обучение»

Актуальность работы не вызывает сомнений, поскольку проведённые исследования направлены на решение современных проблем, к числу которых относятся различные задачи корпусной лингвистики и искусственного интеллекта. Результаты работы имеют потенциал применения в таких отраслях, как машинный перевод и голосовые ассистенты.

В целом полученные научно-практические результаты и практическая значимость работы не вызывают сомнений. Вместе с тем, к тексту автореферата имеются замечания.

1. Основной мотивацией исследования является поиск альтернативы моделям типа «черного ящика» для задач анализа текстов. Однако, сегодня существует ряд методов интерпретации результатов работы подобных моделей (например, SHapley Additive exPlanations (SHAP) Partial Dependence Plots (PDP), Individual Conditional Expectation (ICE) plots, Local Interpretable Model-agnostic Explanations (LIME)). Однако в тексте автореферата такие методы не упоминаются и сравнения с ними не приводятся.

2. В качестве ключевой цели работы сформулирована разработка методов анализа текстов, обеспечивающих интерпретируемость результатов. При этом одной из основных областей приложения результатов в работе являются психолингвистические исследования. Не ясно, чем обусловлен выбор именно такой области приложения, где могут допускаться различные толкования результатов? Почему выбор не сделан в пользу более строгих областей, например, таких как медицина, с необходимостью интерпретации результатов при постановке медицинского диагноза, или как банковского

бизнеса, в котором регулятор требует прозрачности моделей принятия решения?

3. Предлагаемые результаты исследований получены на основе методов, являющихся альтернативными к современным подходам на основе больших языковых моделей. Критикуется недетерминированность и непредсказуемость любого ответа, полученного на основе большой языковой модели. Однако, повышение детерминизма неизбежно приводит к понижению глубины поиска. Известные языковые модели нацелены на поиск приемлемого компромисса между глубиной поиска и предсказуемостью результата. Как именно такой компромисс достигается предлагаемыми решениями из текста автореферата неясно.

4. Для моделей, основанных на векторном представлении слов (например, Word2vec) справедливы замечания соискателя о дефиците анализа связей между словами и трудностях анализа связного текста. Однако, нейросетевые модели, реализованные в архитектуре «Трансформер», взаимосвязь между словами учитывают, и позволяют анализировать семантику текста. Как соотносятся результаты решений соискателя с результатами на подобных архитектурах из текста автореферата не ясно.

Данные замечания имеют рекомендательный характер и не влияют на качество диссертационной работы.

Диссертационная работа Смирнова И.В. соответствует критериям п.9 «Положения о присуждении ученых степеней» ВАК Минобрнауки РФ, утвержденного постановлением Правительства РФ от 24.09.2013 г. № 842, а ее автор, Смирнов Иван Валентинович, заслуживает присуждения ученой степени кандидата технических наук по специальности 1.2.1 – «Искусственный интеллект и машинное обучение».



« 16 » апреля 2024 г.

Директор института прикладной математики и компьютерных наук,  
доктор технических наук, профессор, федеральное государственное  
автономное образовательное учреждение высшего образования  
«Национальный исследовательский Томский государственный  
университет», Российская Федерация, 634050, г. Томск, пр. Ленина, 36,  
(3822)52-94-96, zamyatin@mail.tsu.ru, www.tsu.ru

Замятин Александр Владимирович

Доцент кафедры института прикладной математики и компьютерных наук,  
кандидат технических наук, федеральное государственное автономное  
образовательное учреждение высшего образования «Национальный  
исследовательский Томский государственный университет», Российская  
Федерация, 634050, г. Томск, пр. Ленина, 36, 7(906)956-77-85,  
marigostra@yandex.ru, www.tsu.ru.

Пожидаев Михаил Сергеевич



ПОДПИСЬ УДОСТОВЕРЯЮ  
ВЕДУЩИЙ ДОКУМЕНТОВЕД  
АНДРИЕНКО И. В.