

ОТЗЫВ

на автореферат диссертации Ватолина Алексея Сергеевича на тему «Обучение и оценивание мультязычных нейросетевых моделей семантического векторного представления научных текстов», представленной на соискание ученой степени кандидата технических наук по специальности 2.3.8 «Информатика и информационные процессы».

В условиях значительного роста объемов научной информации, фиксируемого в международных и российских наукометрических базах данных (Scopus, eLibrary.ru), критически важной становится задача автоматизированного анализа и поиска релевантных научных текстов. Существующие поисковые системы, основанные на лексическом совпадении, уже не способны эффективно справляться с семантической сложностью и многоязычностью современного научного дискурса. В связи с этим диссертационное исследование А.С. Ватолина, посвященное разработке специализированных мультязычных нейросетевых моделей для векторизации научных текстов и инструментов их оценки, является весьма актуальным. Особую значимость работе придает фокус на поддержке русского языка, для которого наблюдается дефицит подобных специализированных инструментов.

Научная новизна этого исследования заключается в системном подходе к разработке и оценке векторных моделей, предназначенных специально для научных текстов. Автор предлагает процедуру обучения моделей для векторизации, которая показывает высокое качество работы. При этом модели имеют значительно меньше параметров и работают намного быстрее по сравнению с универсальными моделями, которые применяются в самых разных областях. Также предложенная полуавтоматизированная методика создания RuSciFact с использованием больших языковых моделей и экспертной валидации открывает новые перспективы для генерации сложных проверочных наборов данных.

Практическая значимость работы подтверждается успешным внедрением разработанной модели SciRus-tiny в промышленную эксплуатацию на портале Научной электронной библиотеки eLibrary.ru в качестве основы сервиса «нейропоиск». Это напрямую способствует повышению эффективности работы российских исследователей с научной информацией. Кроме того, интеграция разработанных бенчмарков в международный проект MTEB обеспечивает воспроизводимость результатов и встраивает российские разработки в глобальный научный контекст. Предложенные модели SciRus-small-cite и SciRus-tiny-cite занимают соответственно первое и второе места на русскоязычном подмножестве задач RuSciBench среди рассмотренных моделей, несмотря на существенно меньшее число параметров по сравнению с крупными универсальными моделями.

В качестве рекомендации можно отметить, что целесообразно было бы рассмотреть дополнительные способы оценки качества кросс-языкового выравнивания векторных пространств, например, через анализ устойчивости ближайших соседей при малых возмущениях запросов. Данное замечание носит рекомендательный характер и не снижает общей высокой оценки проведенного исследования.

Диссертационная работа Ватолина Алексея Сергеевича является завершённой самостоятельной научно-квалификационной работой, содержащей предложения по решению важной задачи разработки эффективных методов векторного представления научных текстов для русского и английского языков.

Работа соответствует критериям п. 9 «Положения о присуждении ученых степеней», предъявляемым к диссертациям на соискание ученой степени кандидата наук, а её автор, Ватолин Алексей Сергеевич, заслуживает присуждения искомой ученой степени кандидата технических наук по специальности 2.3.8 – «Информатика и информационные процессы».

**Декан факультета «Информационные технологии»
ФГБОУ ВО «Московский государственный психолого-педагогический университет»,
зав. кафедрой «Прикладная информатика и мультимедийные технологии»,
Почетный работник науки и техники РФ, лауреат премии Правительства РФ в
области образования, доктор технических наук, профессор**

09.09.2021

Лев Семёнович Куравский



П/П заверяю
Заведующая сектором
отдела по работе с персоналом
Завизалкина Е.Ю.

Факультет «Информационные технологии» Московского государственного психолого-педагогического университета, 107143, г. Москва, Открытое шоссе, д.24 к.27

Тел.: 8(499)167-66-74