

ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

на диссертационную работу Скачкова Николая Андреевича

«Вероятностные нейросетевые модели понимания

и генерации текстов естественного языка»,

представленную на соискание ученой степени кандидата технических наук по специальности 2.3.8 «Информатика и информационные процессы»

В условиях стремительного развития методов машинного обучения и широкого внедрения нейросетевых технологий задача высококачественного автоматического машинного перевода сохраняет статус одной из ключевых задач в области искусственного интеллекта. Несмотря на значительный прогресс, современные системы перевода по-прежнему сталкиваются с систематическими ошибками достоверности, гладкости и контекстной согласованности. Данные проблемы особенно заметны при автоматическом переводе длинных текстов, специальной лексики и на малоресурсных языковых направлениях. Диссертационная работа Скачкова Николая Андреевича посвящена комплексному решению этих проблем и предлагает ряд новых моделей и подходов для повышения качества нейросетевого машинного перевода и решения ключевых проблем, возникающих при обучении подобных систем.

Скачков Николай Андреевич начал исследования под моим руководством в 2017 году, распределившись в бакалавриате на кафедру математических методов прогнозирования ВМК МГУ. Уже на третьем курсе получил собственные результаты, которые опубликовал и доложил на конференции по компьютерной лингвистике «Диалог». В магистратуре и аспирантуре демонстрировал высокий уровень самостоятельности, глубокое понимание предметной области и нацеленность на получение практически значимых научных результатов. В ходе работы соискатель проявил себя как зрелый и самостоятельный исследователь, способный ставить и решать сложные научные задачи.

К числу центральных результатов диссертации относятся несколько разработанных методов, позволяющих преодолеть различные ограничения, возникающие при обучении моделей перевода. Так, важным результатом является теоретически обоснованный метод совместного обучения прямой и обратной моделей перевода на основе максимизации циклического правдоподобия. Предложена устойчивая процедура дообучения без привлечения вспомогательных языковых моделей и сложных эвристик; показано улучшение качества перевода и сохранения смысла на малоресурсных направлениях при совместном дообучении с обратной моделью.

Кроме этого, важным результатом является метод переупорядочивания гипотез переводной модели на основе человеческой разметки для устранения систематических ошибок, таких как недостаточные, избыточные и неточные переводы. Подход позволяет эффективно и экономично адаптировать системы к специализированным доменам, в частности, к электронной коммерции, и дает измеримое улучшение качества как по автоматическим метрикам качества, так и с точки зрения независимых экспертов-разметчиков.

Третьим центральным результатом диссертации является интеграция метода маскировки входа в процесс обучения моделей перевода. В рамках этого подхода была выведена новая вероятностная постановка и реализована соответствующая функция потерь, стимулирующая лучшее моделирование контекстных зависимостей в задаче перевода. Показана возможность использования одноязычных данных, а также естественная адаптация к задачам постредактирования. Использование данного подхода показывает прирост качества относительно использования классической авторегрессионной функции потерь.

Последним ключевым результатом работы является сегментация длинных текстов на основе тематического моделирования (ARTM) с учетом сегментной структуры документа. Предложены модификации, улучшающие метрики сегментации и позволяют улучшить качество документного перевода за счет разделения входных документов на семантически однородные блоки. Эксперименты на реальных данных подтверждают важность семантически осмысленной сегментации для получения контекстно согласованного перевода.

Для объективной оценки эффективности предложенных подходов разработана и применена методика комплексного экспериментального анализа, включающая сравнение с сильными базовыми системами по стандартным метрикам с использованием подготовленных профессиональных переводов, а также по нейросетевым метрикам качества и по независимой оценке профессиональных переводчиков. Результаты показывают устойчивые улучшения как в среднем качестве, так и по ключевым типам ошибок, значимые для конечных пользователей.

Основные результаты диссертации Н. А. Скачкова доложены на авторитетных российских конференциях, включая «Ломоносов» (2021), ММРО-2021, «Интеллектуализация обработки информации» (2022, 2024). Работы опубликованы в журналах, рекомендованных ВАК, и периодических изданиях, индексируемых в международных базах Scopus. Это свидетельствует о высокой

научной и практической значимости полученных результатов и признании в профессиональном сообществе.

Диссертационная работа «Вероятностные нейросетевые модели понимания и генерации текстов естественного языка» представляет собой завершённое самостоятельное научное исследование, содержащее новые, теоретически обоснованные и практико-ориентированные результаты в области нейросетевого машинного перевода. Полученные научные положения, методы и экспериментальные подтверждения вносят вклад в развитие вычислительной лингвистики, повышая качество, достоверность, и контекстную согласованность систем автоматического перевода.

Также стоит отметить личные качества Н.А. Скачкова: инициативность, высокая работоспособность, ответственность, умение вести глубокий научный поиск, аккуратность в экспериментальной работе, и нацеленность на результат.

С учетом изложенного считаю, что диссертационная работа Скачкова Николая Андреевича «Вероятностные нейросетевые модели понимания и генерации текстов естественного языка» отличается научной новизной, теоретической и практической значимостью и может быть представлена к защите в диссертационный совет по специальности 2.3.8 «Информатика и информационные процессы».

Научный руководитель:

доктор физико-математических наук, профессор РАН,
заведующий кафедрой математических методов
прогнозирования факультета вычислительной математики и
кибернетики ФГБОУВО «Московский государственный
университет имени М. В. Ломоносова»

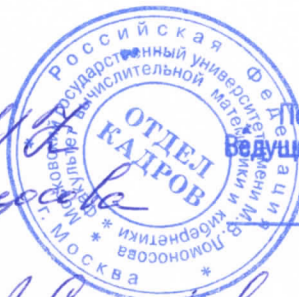
Воронцов Константин Вячеславович
k.vorontsov@iai.msu.ru

Подпись

Дата

18.11.2025

Декан факультета
МТЧ и
академ



Подпись удостоверяю
Будущий специалист по кадрам

Т.Г. Коваленко