Отзыв зарубежного научного руководителя
о работе докторанта кафедры ИСЭ ВШЭБ
Токтоналы Аиды над диссертационным исследованием
на тему «Разработка системы машинного перевода на основе глубокого обучения для перевода терминов IT сферы»

Докторант Токтоналы Аида была зачислена в докторантуру PhD по специальности 710100 «Компьютерные и информационных технологии» образовательной программы «Информационные системы и процессы» в ноябре 2022 года (Приказ №1/54 от 09.11.2022).

Тема диссертационной работы Токтоналы А. «Исследование селевых потоков юга Кыргызстана с использованием технологии машинного обучения» была утверждена Приказом №2/13 от 31 марта 2023 года.

По проекту Erasmus-DERECKA «The Development of Doctoral Education and Research Capacities of Kyrgyzstan Academy» № 609918-EPP-1-2019-1-KG-EPPKA2-CBHE-JP прошла зарубежную научную стажировку в Гамбургском университете прикладных наук (г. Гамбург, Германия) с 20.11.2022 по 24.12.2022.

Стажировка позволила докторанту Токтоналы А. повысить профессиональную квалификацию, уровень методологической и теоретической подготовки для проведения научных исследований.

В процессе сбора данных для исследования по теме «Исследование селевых потоков юга Кыргызстана с использованием технологии машинного обучения» возникли трудности, связанные с ограниченностью доступных данных, что затруднило полноценное применение методов машинного обучения. Посоветовавшись руководителем Орозобековой А. К. и мною 24-января 2024-года докторантом Токтоналы А. было подано заявление на частичное изменение темы диссертации.

Частичное изменение темы позволило продолжить работу, сохраняя основную направленность исследования (применение методов машинного обучения и глубокого обучения).

Докторант Токтоналы А. работает над диссертационной работой по измененной теме «Разработка системы машинного перевода на основе глубокого обучения для перевода терминов IT сферы» (Утверждена приказом № 2/12 от «6» февраля 2024 г.). Подготовила наборы данных для создания терминологической базы для обучения и тренировки модели, по результатам которого опубликовано 1 статья в журнале «Известия КГТУ» (DOI:10.56634-16948335.2024.4.1129-1136).

Докторант Токтоналы А. проделала большую работу по созданию и подготовке параллельного корпуса, который является основой для обучения модели. Далее применяя современные технологии обучения (трансформеры, модели нейронного машинного перевода) можно улучшить качество машинного перевода технических текстов и IT терминов в различных контекстах.

Хотела бы обратить внимание на точность предобработки терминов, чтобы избежать потери значимости слов и фраз в процессе токенизации и нормализации. Было бы полезно расширить параллельный корпус, добавив больше разнообразных примеров с различными вариациями в переводе терминов, что позволяет повысить точность модели.

Рекомендуется более детально оценить результаты работы модели с использованием стандартных метрик качества перевода, таких как BLEU, чтобы оценить эффективность модели в реальных условиях.

В целом, исследовательская работа докторанта Токтоналы А. является значимым вкладом в развитие области машинного перевода и обработки специализированной терминологии. Модель, обученная на параллельном корпусе IT терминов, имеет высокий потенциал для применения в реальных проектах. Уверена, что дальнейшая доработка и расширение корпуса сделает модель еще более эффективной.

Зарубежный научный руководитель:

Prof. Dr.-Ing. Marina Tropmann-Frick

Marina.Tropmann-Frick@haw-hamburg.de

HOCHSCHULE FÜR ANGEWANDTE WISSENSCHAFTEN HAMBURG
Fakultät Technik und Informatik

Department Informatik

HAW HAMBURG