

Письменный отзыв официального рецензента
на диссертационную работу Султановой Назерке Жолдыбаевны
«Open vocabulary model for Kazakh language using deep neural networks», представленную на соискание степени доктора философии (PhD) по специальности 6D070400 – «Вычислительная техника и программное обеспечение»

№ п/п	Критерии	Соответствие критериям	Обоснование позиции официального рецензента
1.	Тема диссертации (на дату ее утверждения) соответствует направлениям развития науки и/или государственным программам	1.1 Соответствие приоритетным направлениям развития науки и/или государственным программам 1) Диссертация выполнена в рамках проекта или целевой программы, финансируемого(ой) из государственного бюджета (указать название и номер проекта или программы) 2) Диссертация выполнена в рамках другой государственной программы (указать название программы) 3) <u>Диссертация соответствует приоритетному направлению развития науки, утвержденному Высшей научно-технической комиссией при Правительстве Республики Казахстан (указать направление)</u>	Тема диссертация посвящена актуальной проблеме в сфере искусственного интеллекта «Open vocabulary model for Kazakh language using deep neural networks» которая соответствует приоритетному направлению специальности 6D070400 – «Вычислительная техника и программное обеспечение».
2.	Важность для науки	Работа вносит/не вносит существенный вклад в науку, а ее важность хорошо раскрыта/не раскрыта	В работе представлены новые данные о моделировании казахского языка с использованием нейронных сетей и механизма внимания. Данная работа вносит существенный вклад в науку об обработке естественного языка и важность работы раскрыта достаточно. Важность работы заключается в разработке новой архитектуры для генерации текста на казахском языке.

3.	Принцип самостоятельности	Уровень самостоятельности: 1) <u>Высокий</u>; 2) Средний; 3) Низкий; 4) Самостоятельности нет	Данная работа состоит из обширного литературного обзора, обработке данных, экспериментов, а также анализа работы в целом. Следовательно, диссертация является самостоятельно выполненной актуальной работой, имеет научную и практическую значимость.
4.	Принцип внутреннего единства	4.1 Обоснование актуальности диссертации: 1) <u>Обоснована</u>; 2) Частично обоснована; 3) Не обоснована.	Для решения поставленных задач в диссертации представлены и обоснованы новые технологии искусственного интеллекта. Достоверность и обоснованность научных результатов не вызывает сомнений, так как применялись апробированные методы для тренировки модели и выявления результатов.
		4.2 Содержание диссертации отражает тему диссертации: 1) <u>Отражает</u>; 2) Частично отражает; 3) Не отражает	Содержание диссертации отражает тему диссертации.
		4.3. Цель и задачи соответствуют теме диссертации: 1) <u>соответствуют</u>; 2) частично соответствуют; 3) не соответствуют	Цели и задачи представленные в диссертационной работе полностью соответствуют теме диссертации.
		4.4 Все разделы и положения диссертации логически взаимосвязаны: 1) <u>полностью взаимосвязаны</u>; 2) взаимосвязь частичная; 3) взаимосвязь отсутствует	Все разделы и положения диссертации логически полностью взаимосвязаны, содержат совокупность новых научных результатов, обладают внутренним единством и целостностью.
		4.5 Предложенные автором новые решения (принципы, методы) аргументированы и оценены по сравнению с известными решениями:	Диссертантом в полном объеме рассмотрены известные способы моделирования языка как традиционные, так и с использованием нейронных

		1) критический анализ есть; 2) анализ частичный; 3) анализ представляет собой не собственные мнения, а цитаты других авторов	сетей. А также на основании представленных методов была разработана модель для генерации казахского языка. Тем не менее, недостаточно полно представлены результаты отечественных исследователей для критического анализа.
5.	Принцип научной новизны	5.1 Научные результаты и положения являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%) 5.2 Выводы диссертации являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%) 5.3 Технические, технологические, экономические или управленческие решения являются новыми и обоснованными? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%)	Научные результаты представленные в диссертационной работе несомненно имеют теоретическую и практическую ценность, тем не менее научные результаты и положения являются частично новыми. Выводы и результаты диссертации основаны на большом количестве экспериментов и литературном анализе. Выводы и полученные результаты имеют достаточно новый и подтвержденный материал. Технические, технологические, экономические или управленческие решения являются частично новыми и обоснованными. Полученные результаты могут быть использованы на производстве.
6.	Обоснованность основных выводов	Все основные выводы <u>основаны</u>/не основаны на весомых с научной точки зрения доказательствах либо достаточно хорошо обоснованы (для qualitative research и направлений подготовки по искусству и гуманитарным наукам)	Достоверность полученных результатов обоснована применением диссертантом комплекса современных методов исследования и анализа, соответственно все основные выводы диссертационной работы основаны на весомых с научной точки зрения доказательствах, имеют основание на предыдущих результатах исследований.
7.	Основные положения, выносимые на защиту	Необходимо ответить на следующие вопросы по каждому положению в отдельности: 7.1 Доказано ли положение? 1) <u>доказано</u>;	Все основные положения выносимые на защиту полностью доказаны.

		2) скорее доказано; 3) скорее не доказано; 4) не доказано	
		7.2 Является ли тривиальным? 1) <u>да</u>; 2) <u>нет</u>	Основные положения, выносимые на защиту, не являются тривиальными так как содержат решения, отличающиеся научной новизной и теоретической ценностью.
		7.3 Является ли новым? 1) <u>да</u>; 2) <u>нет</u>	Основные положения диссертационной работы выносимые на защиту являются новыми.
		7.4 Уровень для применения: 1) узкий; 2) средний; 3) <u>широкий</u>	Уровень применения языковой модели представленной в диссертационной работе – широкий. Модель применима не только для обработки естественного языка, а также в области компьютерного зрения и голосового распознавания.
		7.5 Доказано ли в статье? 1) <u>да</u>; 2) <u>нет</u>	В рамках диссертационной работы подготовлено и опубликовано 11 научных работ по рассматриваемой теме, из которых: 3 статьи в зарубежном издании и в международных рецензируемых научных журналах; 4 статьи опубликованы в издательствах, соответствующих требованиям высшей аттестационной комиссии Министерства образования и науки Республики Казахстан; 4 статьи опубликованы в трудах международных конференций.
8.	Принцип достоверности Достоверность источников и предоставляемой информации	8.1 Выбор методологии - обоснован или методология достаточно подробно описана 1) <u>да</u>; 2) <u>нет</u>	Основными исследовательскими подходами являются методы вычислительного интеллекта. Глубокие нейронные сети применялись для построения генеративной языковой модели для казахского языка. Применяемые методы и модели полностью описаны.

	8.2 Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий: 1) <u>да</u> ; 2) <u>нет</u>	Для работы с данными была использована библиотека Python nltk. Нейронная модель была реализована при помощи библиотеки tensorflow. Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий.	
	8.3 Теоретические выводы, модели, выявленные взаимосвязи и закономерности доказаны и подтверждены экспериментальным исследованием: 1) <u>да</u> ; 2) <u>нет</u>	Теоретические выводы, модели, выявленные взаимосвязи и закономерности доказаны и подтверждены экспериментальным исследованием. Однако, не было уделено достаточного внимания механизму внимания примененного для казахского языка.	
	8.4 Важные утверждения <u>подтверждены</u> /частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу	Все важные утверждения в данной работе подтверждены ссылками на актуальную и достоверную научную литературу.	
	8.5 Использованные источники литературы <u>достаточно</u> /не достаточно для литературного обзора	Автором диссертации сделан достаточный литературный обзор научной литературы по теме исследования, использовано 108 научных работ для литературного обзора.	
9.	Принцип практической ценности	9.1 Диссертация имеет теоретическое значение: 1) <u>да</u> ; 2) <u>нет</u>	Целью диссертации является разработка модели для генерации текста. Инновационная архитектура и система была разработана соискателем. Из этого следует что диссертация имеет теоретическое значение.
		9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике: 1) <u>да</u> ; 2) <u>нет</u>	Результаты приведенные в диссертационной работе были опубликованы в международных журналах и научно практических конференциях. Результаты работы имеют практическую ценность для применения в индустрии.
		9.3 Предложения для практики являются новыми? 1) <u>полностью новые</u> ;	Научная новизна темы исследования определяется тем, что построена инновационная языковая модель.

		2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%)	Существенная новизна отличается от предыдущих моделей, основанных на применении нейронных сетей с использованием графического процессора, что делает вычисления более эффективными. Построенная архитектура со слоем внимания повысила производительность модели. Можно заключить что степень новизны практических предложений и рекомендаций высокая.
10.	Качество написания и оформления	Качество академического письма: 1) высокое; 2) среднее; 3) ниже среднего; 4) низкое.	Качество академического письма – высокое, общее оформление работы соответствует требованиям, предъявляемым к докторским диссертациям.

Заключение:

Представленная диссертация Султановой Назерке Жолдыбаевны на тему «Open vocabulary model for Kazakh language using deep neural networks» отвечает всем требованиям Правил присуждения степеней, предъявляемым к докторским (PhD) диссертациям и рекомендуется к защите на соискание степени доктора философии (PhD) по специальности 6D070400 – «Вычислительная техника и программное обеспечение».

Официальный рецензент:

PhD, к.т.н., профессор кафедры «Программная инженерия»
НАО Казахский национальный исследовательский
технический университет имени К.И. Сатпаева

14.12.2021г.

Еримбетова А.С.

