

**Письменный отзыв официального рецензента
на диссертационную работу
Оралбековой Дины Орымбаевны**

**на тему «Разработка системы автоматического распознавания речи на основе интегрального подхода»,
представленную на соискание степени доктора PhD по образовательной программе 8D06103 – «Management
information systems»**

№п/п	Критерии	Соответствие критериям (необходимо отметить один из вариантов ответа)	Обоснование позиции официального рецензента
1.	Тема диссертации (на дату ее утверждения) соответствует направлениям развития науки и/или государственным программам	<p>1.1 Соответствие приоритетным направлениям развития науки или государственным программам:</p> <p>1) Диссертация выполнена в рамках проекта или целевой программы, финансируемого(ой) из государственного бюджета (указать название и номер проекта или программы)</p> <p>2) Диссертация выполнена в рамках другой государственной программы (указать название программы)</p> <p>3) <u>Диссертация соответствует приоритетному направлению развития науки, утвержденному Высшей научно-технической комиссией при Правительстве Республики Казахстан (указать направление)</u></p>	<p>Тема диссертации соответствует приоритетному направлению развития науки, утвержденным Высшей научно-технической комиссией при Правительстве Республики Казахстан: «Информационные, коммуникационные и космические технологии».</p>
2.	Важность для науки	Работа вносит/не вносит существенный вклад в науку, а ее важность хорошо раскрыта/не раскрыта	<p>Работа вносит существенный вклад в решение важной социально-значимой проблемы создания средств быстрого и качественного распознавания казахского языка.</p> <p>Основные результаты диссертационной работы по применению новых методов и подходов исследования по направлению распознавания речи могут значительно повлиять на развитие области компьютерной лингвистики, улучшить исследование в области речевого анализа и фонетики казахского языка,</p>

<p>развитие машинного обучения и искусственного интеллекта, математических моделей и т.д.</p> <p>Важность работы заключается в том, что разработанная система автоматического распознавания речи учитывает морфологическое значение слова при декодировании, что дает возможность лучше распознать схожие окончания казахских слов. Реализованная модель по сравнению с другими интегральными моделями, учитывает не только текущие выходы и входы, но и предыдущие выходы и входы для декодирования выходной последовательности.</p> <p>В качестве замечания можно указать следующее. В работе приводится подробное описание реализации аппаратной части распознавания, данная информация представляет практический интерес, но в рамках диссертационной работы можно было бы сократить её объём.</p>			
<p>Диссертационная работа Оралбековой Д.О. представляет новое, самостоятельное и имеющее практическую и научную значимость исследование, свидетельствует о личном вкладе автора диссертации в науку.</p> <p>В работе использованы новые подходы в разработке системы автоматического распознавания речи для языков с ограниченными обучающими данными.</p>	<p>Уровень самостоятельности:</p> <ol style="list-style-type: none"> 1) <u>Высокий</u>; 2) <u>Средний</u>; 3) <u>Низкий</u>; 4) Самостоятельности нет 	<p>Принцип самостоятельности и</p>	<p>3.</p>
<p>Актуальность диссертации заключается в разработке эффективных моделей и алгоритмов для повышения качества распознавания, автоматического декодирования слитной разговорной речи, создания лучшего речевого корпуса для малоресурсных языков, в частности для казахского языка.</p>	<p>4.1 Обоснование актуальности диссертации:</p> <ol style="list-style-type: none"> 1) <u>Обоснована</u>; 2) <u>Частично обоснована</u>; 3) <u>Не обоснована</u>. 	<p>Принцип внутреннего единства</p>	<p>4.</p>
<p>Содержание диссертации в полном объеме отражает цель и задачи исследования.</p>	<p>4.2 Содержание диссертации отражает тему диссертации:</p> <ol style="list-style-type: none"> 1) <u>Отражает</u>; 2) <u>Частично отражает</u>; 3) <u>Не отражает</u> 		

	<p>4.3. Цель и задачи соответствуют теме диссертации: 1) <u>соответствуют</u>; 2) <u>частично соответствуют</u>; 3) <u>не соответствуют</u></p> <p>4.4 Все разделы и положения диссертации логически взаимосвязаны: 1) <u>полностью взаимосвязаны</u>; 2) <u>взаимосвязь частичная</u>; 3) <u>взаимосвязь отсутствует</u></p>	<p>Поставленные в работе цель и задачи исследования полностью соответствуют теме диссертации и раскрывают ее содержание.</p> <p>Диссертация состоит из введения, четырех глав, заключения. В разделе введения автор описывает актуальность исследования, поставленные задачи, новизну. В дальнейших разделах автор описывает работу по сбору и расширению речевого корпуса казахского языка, а также приведена методика по транскрибированию телефонных/мобильных разговоров. В отдельных разделах соискатель приводит полное описание модели кодер-декодер с механизмом внимания и выделяет некоторые улучшения и модификацию для повышения качества распознавания речи. В последнем разделе автор приводит реализацию программного и аппаратного обеспечения интегральной модели распознавания казахской речи. В качестве замечания можно указать следующее. В диссертации разработана модель на основе кодер-декодера с вниманием. Соискателю следовало бы подробнее описать то, как данная модель превосходит другие интегральные модели при распознавании речи на казахском языке. В целом, можно прийти к выводу, что работа обладает внутренним единством полученных результатов и взаимосвязанностью по исследуемым вопросам.</p>
	<p>4.5 Предложенные автором новые решения (принципы, методы) аргументированы и оценены по сравнению с известными решениями: 1) <u>критический анализ есть</u>; 2) <u>анализ частичный</u>; 3) <u>анализ представляет собой не собственные мнения, а цитаты других авторов</u></p>	<p>В работе приведен подробный сравнительный анализ существующих методов распознавания речи на основе интегральной архитектуры, обосновано применение моделей с механизмом внимания для увеличения эффективности работы систем распознавания. Доказана эффективность разработанной системы в ходе тестирования моделей с вниманием.</p>

5.	Принцип научной новизны	<p>5.1 Научные результаты и положения являются новыми?</p> <ol style="list-style-type: none"> 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%) 	<p>Научная новизна и практическая значимость данной диссертационной работы являются принципиально новыми в части:</p> <ul style="list-style-type: none"> - разработки речевого и текстового корпусов для казахского языка; - разработки интегральной модели с модифицированным механизмом внимания для распознавания казахской речи, которая дает возможность корректно распознавать схожие окончания казахских слов в речи, т.к. казахский язык имеет агглютинативный характер; - разработки эффективного алгоритма для распознавания казахской речи на основе модели кодер-декодер; - разработки программного обеспечения, которое автоматически преобразует речь в текст в режиме реального времени. <p>Новизна полученных результатов стала основанием для публикации результатов работы в высокорейтинговых журналах, включая международные научные издания, имеющие ненулевой импакт-фактор и включенные в информационные базы данных Scopus и WoS.</p> <p>В качестве замечания можно указать следующее: при проведении практического эксперимента необходимо было привести больше информации о блоке выделения признаков.</p>
		<p>5.2 Выводы диссертации являются новыми?</p> <ol style="list-style-type: none"> 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%) 	<p>Полученные результаты и выводы базируются на большом экспериментальном и теоретическом материале. Выводы и полученные научные результаты представляют новый материал, доказательно подтвержденный результатами детального анализа экспериментальных данных, полученных в ходе исследований моделей и методов на основе интегрального подхода.</p>
		<p>5.3 Технические, технологические, экономические или управленческие решения являются новыми и обоснованными:</p> <ol style="list-style-type: none"> 1) полностью новые; 	<p>Результатом диссертационной работы является система автоматического распознавания речи, которая разделяет речь, когда говорят 2 человека и автоматически преобразует ее в текст в отдельные блоки. Полученные результаты могут быть</p>

		<p>2) <u>частично новые</u> (новыми являются 25-75%); 3) <u>не новые</u> (новыми являются менее 25%)</p>	<p>использованы для голосового машинного перевода, для компьютерного стенографирования, для голосового управления компьютером, робототехническими и автоматизированными системами, что позволит людям с ограниченными возможностями одновременно осуществлять несколько функций, не связанных с устройствами ввода в машину.</p>
6.	Обоснованность основных выводов	<p>Все основные выводы основаны/не основаны на весомых с научной точки зрения доказательствах либо достаточно хорошо обоснованы (для qualitative research и направлений подготовки по искусству и гуманитарным наукам)</p>	<p>Выводы, приведенные в заключении сделаны на основе проведенных теоретических и практических исследований и полученных экспериментальных результатов. Они полностью подтвердили положения диссертационной разработки.</p>
7.	Основные положения, выносимые на защиту	<p>Необходимо ответить на следующие вопросы по каждому положению в отдельности: 7.1 Доказано ли положение? 1) <u>доказано</u>; 2) скорее доказано; 3) скорее не доказано; 4) не доказано 7.2 Является ли тривиальным? 1) да; 2) <u>нет</u> 7.3 Является ли новым? 1) <u>да</u>; 2) нет 7.4 Уровень для применения: 1) узкий; 2) средний; 3) <u>широкий</u> 7.5 Доказано ли в статье? 1) <u>да</u>; 2) нет</p>	<p>Представленные основные положения, выносимые на защиту доказаны экспериментально и подтверждены результатами выполненных теоретических и практических экспериментов. Элементы тривиальности в данной диссертационной работе отсутствуют. Все найденные закономерности и особенности изученных методов и моделей рассматривались не упрощенно, а с позиции современных знаний в области информационных систем. Положения, выносимые на защиту, диссертанткой являются новыми, так как ранее подобные положения и результаты исследований по теме диссертации не описаны в литературе. Положения, выносимые на защиту имеют широкий уровень применения, так как могут быть использованы при синтезе речи, аннотировании баз данных, а также при автоматическом синхронном переводе с казахского языка на другой язык, идентификации речи и других областях знаний. Разработанная модель распознавания речи может быть применена в робототехнике, системах управления оборудованием и средствах коммуникаций. По результатам диссертационных исследований опубликованы 13 статей, из них 4 в наукометрических базах Scopus и WoS с (1 статья с процентилем 91% Q1, 1 статья с процентилем 93%</p>

		<p>Q2 и 2 статьи с процентилем 43% Q3), 3 статьи в изданиях рекомендуемых КОКсон, 6 статьи в международных научных конференциях, были получены 3 авторских свидетельства на программные обеспечения и 1 патент на изобретение.</p>
<p>8. Принцип достоверности Достоверность источников и предоставляемой информации</p>	<p>8.1 Выбор методологии - обоснован или методология достаточно подробно описана 1) да; 2) нет</p> <p>8.2 Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий: 1) да; 2) нет</p>	<p>Выбор методологии диссертантом сделан на основании литературного анализа методов и моделей распознавания речи на основе интегрального подхода. Также дано полное описание структуры данных корпусов и их особенности.</p>
	<p>8.3 Теоретические выводы, модели, выявленные взаимосвязи и закономерности доказаны и подтверждены экспериментальным исследованием (для направлений подготовки по педагогическим наукам результаты доказаны на основе педагогического эксперимента): 1) да; 2) нет</p>	<p>Для реализации поставленных целей и задач исследования в работе в качестве среды разработки был выбран Anaconda Navigator/Jupyter Notebook с языком программирования Python 3.8 и его библиотека Tensorflow. Для тестирования качества распознавания разработанный программный продукт был интегрирован в микрофон SmartMike Duo от компаний Philips.</p>
	<p>8.4 Важные утверждения подтверждены/частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу</p>	<p>В диссертационной работе в полной мере обоснованы теоретические выводы, разработаны в ходе проведения исследований эффективные модель и алгоритм распознавания речи. Разработанная модель кодер-декодер с вниманием прошла процедуру проверки распознавания речи разных дикторов, проведено тестирование в компании Philips, которые показали вполне удовлетворительные результаты, что подтверждается Актом.</p>
	<p>8.5 Использованные источники литературы достаточно/не достаточно для литературного обзора</p>	<p>Важные утверждения подтверждены ссылками на актуальную и достоверную научную литературу.</p> <p>Список литературных источников диссертации насчитывает 111 научных работ, в том числе для литературного обзора 102, которых достаточно для проведения аналитического литературного обзора по теме диссертации. Стоит отметить, что использованы работы, которые касаются не только</p>

интегральных моделей, но и гибридных моделей распознавания.		
<p>В диссертационной работе проведен детальный анализ современных методов и моделей распознавания речи на основе гибридных, традиционных и интегральных систем, разработан эффективный алгоритм, который сокращает время обучения и время декодирования в процессе распознавания. Модель с механизмом внимания позволит избежать излишних вычислительных операций при построении системы распознавания речи.</p>	<p>9.1 Диссертация имеет теоретическое значение: 1) да; 2) нет</p>	Принцип практической ценности
<p>С каждым годом возрастает потребность со стороны пользователей мобильных устройств и стационарных компьютеров в обеспечении лучшего качества и производительности современных речевых технологий. Поэтому необходимо развивать технологии, которые обеспечивают более естественное взаимодействие пользователя с вычислительными и телекоммуникационными комплексами по сравнению со стандартным графическим интерфейсом. С появлением глубокого обучения, распространение получил интегральный метод распознавания, который представляет систему одной нейронной сетью, реализующую прямое отображение входных аудио данных в последовательности слов. Реализация интегрального метода не требует много памяти компьютера и не требует больших вычислений на серверах. В связи с этим, внедрение таких технологий в повседневную жизнь людей особенно актуально, так как их можно использовать локально в мобильных устройствах.</p>	<p>9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике: 1) да; 2) нет</p>	
<p>Практическая значимость полученных результатов предусматривает применение разработанных алгоритмов и программного обеспечения для дальнейшего использования в развитии других технологий, как синтез речи, машинный перевод, голосовая аутентификация и идентификация и т.д. Разработанная система автоматического распознавания</p>	<p>9.3 Предложения для практики являются новыми? 1) <u>полностью новые</u>; 2) <u>частично новые</u> (новыми являются 25-75%); 3) <u>не новые</u> (новыми являются менее 25%)</p>	

		казахской речи может быть внедрена в государственных структурах, ответственных за расширение области применения национальных языков на базе информационных технологий; в мобильных устройствах, в банках, и в сектор производства различных устройств с поддержкой голосовых функций.
10.	<p>Качество академического письма:</p> <p>1) <u>высокое</u>;</p> <p>2) среднее;</p> <p>3) ниже среднего;</p> <p>4) низкое.</p>	<p>Диссертация написана грамотным научно-техническим, доступным для читателя языком. Стилль изложения лаконичный. Формулировки основных положений и выводов носят логический, законченный характер. В работе имеются незначительные стилистические ошибки.</p>

Заключение

Указанные замечания не снижают в целом положительной оценки диссертационной работы. Считаю, что рецензируемая диссертационная работа Оралбековой Дины Орымбаевны, тема: «Разработка системы автоматического распознавания речи на основе интегрального подхода» по своей актуальности, научной новизне, важности для теории и практики, объёму экспериментальных исследований полностью соответствует требованиям, которые предъявляются к диссертациям на соискание степеней PhD Комитетом по контролю в сфере образования и науки МОН РК, а ее автор Оралбекова Дина Орымбаевна заслуживает присуждения степени доктора философии (PhD) по образовательной программе 8D06103 - «Management information systems»

**Рецензент, к.ф.-м. наук, ассоц. профессор
зав. кафедрой «Программная инженерия»**

Подпись заверяю
Начальник отдела кадров

«01» декабря_2022 г.

Молдагулова Айман Николаевна

А.К. Бейсова

