

**Письменный отзыв официального рецензента
на диссертационную работу ОРАЛБЕКОВА ДИНА ОРЫМБАЕВНА
“Разработка системы автоматического распознавания речи на основе интегрального подхода” представленной на соискание степени доктора
философии (PhD)
по специальности 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) <u>Диссертация выполнена в рамках проекта или целевой программы, финансируемого(ой) из государственного бюджета (указать название и номер проекта или программы)</u></p> <p>2) Диссертация выполнена в рамках другой государственной программы (указать название программы)</p> <p>3) Диссертация соответствует приоритетному направлению развития науки, утвержденному Высшей научно-технической комиссией при Правительстве Республики Казахстан (указать направление)</p>	<p>Диссертационная работа выполнена в рамках грантов МОН РК по приоритетному направлению развития науки «Распознавание речи и компьютерная лингвистика» по проектам ГФ: 1) «Разработка технологии мультязычного автоматического распознавания речи с использованием глубоких нейронных сетей» (2018-2020, государственный регистрационный номер: 0118РК00139) 2) «Разработка интегральной (end-to-end) системы автоматического распознавания речи для агглютинативных языков» (2020-2022, государственный регистрационный номер: 0120РК00344) в Институте информационных и вычислительных технологий КН МОН РК.</p>
2.	Важность для науки	Работа вносит /не вносит существенный вклад в науку, а ее важность хорошо раскрыта /не раскрыта	<p>Диссертационная работа вносит существенный вклад в науку в связи с тем, что:</p> <ul style="list-style-type: none"> - разработан корпус в объеме 2000 часов речи с транскрипциями и учетом различных видов речи: подготовленная (чтение), спонтанная; - разработана интегральная модель с применением механизма внимания для

			<p>распознавания казахской речи, отличающаяся интеграцией в архитектуре LSTM, BLSTM и механизма внимания;</p> <p>- разработан алгоритм для эффективного распознавания казахской речи на основе интегрального модуля.</p>
3.	Принцип самостоятельности	<p>Уровень самостоятельности:</p> <p>1) Высокий;</p> <p>2) Средний;</p> <p>3) Низкий;</p> <p>4) Самостоятельности нет</p>	<p>Диссертационная работа Оралбековой Дины Орымбаевной является самостоятельно выполненной автором работой. Разработана и реализована интегральную модель и алгоритм для распознавания казахской речи. Разработан и расширен корпус речи для казахского языка. Выполнена экспериментальная оценка разработанной модели и алгоритма распознавания казахской речи.</p>
4.	Принцип внутреннего единства	<p>4.1 Обоснование актуальности диссертации:</p> <p>1) Обоснована;</p> <p>2) Частично обоснована;</p> <p>3) Не обоснована.</p>	<p>Актуальность диссертации обоснована. Проблема распознавания речи является одной из актуальных направлений искусственного интеллекта. Одной из актуальных проблем в этом направлении является проблема распознавания речи для малоресурсных языков, к которым относится и казахский язык. В истории исследования проблемы распознавания речи существовали различные подходы к данной проблеме и наиболее существенные результаты появились с использованием технологии нейронных сетей глубокого обучения. Однако, применение технологии нейронных сетей глубокого обучения для нового языка требует создания корпуса обучения, построение модели нейронных сетей</p>

			<p>глубокого обучения, дающей максимально хороший результат, что требует значительных исследований. В данной диссертационной работе проведены такие исследования для казахского языка.</p>
		<p>4.2 Содержание диссертации отражает тему диссертации: 1) Отражает; 2) Частично отражает; 3) Не отражает</p>	<p>Содержание диссертации отражает тему диссертации. Диссертация состоит из 4 глав и 5 приложений. Первая глава описывает современные методы автоматического распознавания речи. Вторая глава посвящена разработке речевого корпуса для интегрального распознавания казахской речи. Третья глава посвящена разработке интегральной модели для распознавания казахской речи. Четвертая глава описывает реализацию интегральной системы распознавания казахской речи.</p>
		<p>4.3. Цель и задачи соответствуют теме диссертации: 1) соответствуют; 2) частично соответствуют; 3) не соответствуют</p>	<p>Основной целью исследования являются разработка модели, метода и алгоритма для повышения точности распознавания казахской речи на основе интегральной архитектуры. Цель и задачи полностью соответствуют теме диссертации.</p>
		<p>4.4 Все разделы и положения диссертации логически взаимосвязаны: 1) полностью взаимосвязаны; 2) взаимосвязь частичная; 3) взаимосвязь отсутствует</p>	<p>В диссертации все разделы и их положения полностью взаимосвязаны и имеют достаточную степень логического сопряжения. В диссертации последовательно описана разработка распознавания казахской речи от сбора данных до построение нейронных сетей для тренировки модели.</p>

		<p>4.5 Предложенные автором новые решения (принципы, методы) аргументированы и оценены по сравнению с известными решениями:</p> <p>1) критический анализ есть;</p> <p>2) анализ частичный;</p> <p>3) анализ представляет собой не собственные мнения, а цитаты других авторов</p>	<p>Предложенные автором новые решения аргументированы, критически проанализированы и оценены по сравнению с известными решениями.</p>
5.	Принцип научной новизны	<p>5.1 Научные результаты и положения являются новыми?</p> <p>1) полностью новые;</p> <p>2) частично новые (новыми являются 25-75%);</p> <p>3) не новые (новыми являются менее 25%)</p>	<p>Разработанный обучающий речевой корпус для казахского языка собран и обработан впервые. Разработанная интегральная модель в архитектуре LSTM, BLSTM и механизма внимания для распознавания казахской речи представлена впервые.</p> <p>Новизна разработанного алгоритма для эффективного распознавания казахской речи на основе интегрального модуля заключается в новизне реализации предложенной модели.</p>
		<p>5.2 Выводы диссертации являются новыми?</p> <p>1) полностью новые;</p> <p>2) частично новые (новыми являются 25-75%);</p> <p>3) не новые (новыми являются менее 25%)</p>	<p>Приведенные в диссертации выводы относительно системы автоматического распознавания речи на основе интегрального подхода для казахского языка являются новыми и ранее не были опубликованы.</p>
		<p>5.3 Технические, технологические, экономические или управленческие решения являются новыми и обоснованными:</p> <p>1) полностью новые;</p> <p>2) частично новые (новыми являются 25-75%);</p> <p>3) не новые (новыми являются менее 25%)</p>	<p>Техническая новизна работы заключается в программной реализации системы распознавания казахской речи на основе предложенной интегральной модели распознавания речи.</p>
6.	Обоснованность основных выводов	<p>Все основные выводы основаны/не основаны на весомых с научной точки зрения доказательствах либо достаточно хорошо</p>	<p>Достаточную степень обоснованности результата по разработанному корпусу казахской речи подтверждает анализ научных</p>

		<p>обоснованы (для qualitative research и направлений подготовки по искусству и гуманитарным наукам)</p>	<p>публикации по корпусам речи казахского языка, а также описанные результаты сбора речевых данных по казахскому языку.</p> <p>Степень обоснованности второго результата по разработанной интегральной модели показывает процесс создания и результаты экспериментов интегральной модели LSTM, BLSTM и механизма внимания для распознавания речи на корпусе казахского языка.</p> <p>Третий результат по алгоритму эффективного распознавания казахской речи обоснован результатами экспериментов обучения интегральной модели распознавания модели на различных размерах корпусов казахского языка.</p>
7.	<p>Основные положения, выносимые на защиту</p>	<p>Необходимо ответить на следующие вопросы по каждому положению в отдельности:</p> <p>7.1 Доказано ли положение?</p> <p>1) доказано;</p> <p>2) скорее доказано;</p> <p>3) скорее не доказано;</p> <p>4) не доказано</p> <p>7.2 Является ли тривиальным?</p> <p>1) да;</p> <p>2) нет</p> <p>7.3 Является ли новым?</p> <p>1) да;</p> <p>2) нет</p> <p>7.4 Уровень для применения:</p> <p>1) узкий;</p> <p>2) средний;</p>	<p>Основное положение выносимые автором на защиту, являются новым, не тривиальным и были доказаны и подтверждены научными результатами исследования. На основании полученных экспериментальных данных были опубликованы 3 авторских свидетельства, 1 патент на изобретение и опубликовано 7 работ, из которых 3 статьи опубликованы в журналах, рекомендованных Комитетом по контролю в сфере образования и науки МОН РК, 4 статьи опубликованы в изданиях имеющие ненулевой импакт-фактор, индексируемых базой Scopus. Уровень для применения результатов диссертационной работы является широким.</p>

		<p>3) <u>широкий</u></p> <p>7.5 Доказано ли в статье?</p> <p>1) <u>да</u>;</p> <p>2) нет</p>	
8.	<p>Принцип достоверности</p> <p>Достоверность источников и предоставляемой информации</p>	<p>8.1 Выбор методологии - обоснован или методология достаточно подробно описана</p> <p>1) <u>да</u>;</p> <p>2) нет</p>	<p>Методология исследования достаточно подробно описана. Для обучения сети был разработан корпус в объеме 2000 часов речи с транскрипциями. При создании корпуса учтены различные виды речи: подготовленная (чтение), спонтанная. Разработана интегральная архитектура кодер-декодера с механизмом внимания с помощью нейронных сетей LSTM и BLSTM и обучена на разработанном корпусе речи. Выносимые на защиту положения доказаны экспериментальными результатами.</p>
		<p>8.2 Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий:</p> <p>1) <u>да</u>;</p> <p>2) нет</p>	<p>Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий. Программа была реализована на языке Python. Для проведения вычислений была применена библиотека программного обеспечения с открытым исходным кодом Tensorflow.</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):</p> <p>1) <u>да</u>;</p>	<p>Теоретические выводы, модели, выявленные взаимосвязи и закономерности доказаны и подтверждены экспериментальным исследованием.</p>

		2) <u>нет</u>	
		8.4 Важные утверждения подтверждены /частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу	Важные утверждения подтверждены ссылками на актуальную и достоверную научную литературу.
		8.5 Используемые источники литературы достаточны /не достаточны для литературного обзора	Список использованных источников диссертации насчитывает 111 научных работ, в том числе для литературного обзора 102, которых достаточно для проведения аналитического литературного обзора по теме диссертации.
9	Принцип практической ценности	9.1 Диссертация имеет теоретическое значение: 1) да ; 2) <u>нет</u>	Теоретическая значимость полученных научных результатов заключается в построении и испытании интегральной модели (архитектуры) распознавания казахской речи комбинации LSTM, BLSTM и механизма внимания.
		9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике: 1) да ; 2) <u>нет</u>	Практическая значимость полученных научных результатов заключается в возможности применения разработанных алгоритмов и программного обеспечения, корпуса речи казахского языка для дальнейшего использования в развитии таких технологий, как синтез речи, машинный перевод, голосовая аутентификация и идентификация.
		9.3 Предложения для практики являются новыми? 1) полностью новые ; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%)	Предложения для практики являются новыми, так как разработанная модель распознавания казахской речи комбинации LSTM, BLSTM и механизма внимания обучена на большом объеме речевого корпуса казахского языка в 2000 часов и имеет высокое качество распознавания казахской речи.

