ОТЗЫВ

официального рецензента на диссертационную работу

Ерқос Айнұр на тему «Разработка метода и программного пакета для обработки больших медицинских данных», предоставленную на соискание степени доктора философии (PhD) по специальности «8D06102 – Компьютерная инженерия»

№ п/п	Критерии	Соответствие критериям (подчеркнуть один из вариантов ответа)	Обоснование позиции официального рецензента (замечания выделить курсивом)
1.	Тема диссертации (на дату ее утверждения) соответствует направлениям развития науки и/или государственным программам	1.1 Соответствие приоритетным направлениям развития науки 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) диссертация соответствует приоритетному направлению развития науки, утвержденному Высшей научно-технической комиссией при Правительстве Республики Казахстан (указать направление).	Диссертация выполнена в рамках научных проектов, финансируемых из государственного бюджета: «Персональное прогнозирование выживаемости трансплантата у реципиентов трансплантата почки перед трансплантацией с применением методов машинного обучения» (№ гос. регистрации 0122РК00958, 2022–2024 гг.) и «Разработка алгоритмического механизма для интеллектуально-программного модуля, использующего искусственный интеллект для обработки и анализа данных, связанных с трансплантацией почки» (№ гос. регистрации 0124РК00417, 2024–2026 гг.). По содержанию и направленности диссертация соответствует приоритету «Передовое производство, цифровые и космические технологии».
2.	Важность для науки	Работа вносит/не вносит существенный вклад в науку, а ее важность хорошо раскрыта/не раскрыта.	Диссертационное исследование посвящено актуальной научно-технической задаче обработки несбалансированных медицинских данных и применению методов машинного обучения в трансплантологии. Научная значимость работы определяется разработкой метода NAUS для балансировки шумных и несбалансированных выборок, а также созданием пяти прогностических моделей для оценки клинического состояния после трансплантации, включая прогноз

			жизнеспособности трансплантата на разных временных горизонтах. Интеграция этих моделей в программный пакет усиливает прикладную направленность исследования.
3.	Принцип самостоятельности	Уровень самостоятельности:	Диссертация демонстрирует высокий уровень самостоятельности автора. В работе последовательно представлены аналитический обзор литературы, обоснование метода NAUS, построение пяти моделей машинного обучения, архитектурное описание программного пакета и результаты экспериментов.
		1) высокий;	
		2) средний;	
		3) низкий;	
		4) самостоятельности нет.	
4.	Принцип внутреннего единства	4.1 Обоснование актуальности диссертации:	Актуальность диссертации убедительно обоснована ростом объема и сложности медицинских данных, а также необходимостью более точного прогнозирования исходов почечной трансплантации. Традиционные статистические подходы не в полной мере учитывают нелинейные зависимости и редкие неблагоприятные события, тогда как использование методов машинного обучения для несбалансированных данных является научно обоснованным и востребованным направлением.
		1) обоснована;	
		2) частично обоснована;	
		3) не обоснована.	
		4.2 Содержание диссертации отражает тему диссертации:	Содержание диссертации полностью соответствует заявленной теме. Во введении определены цель, задачи, объект и предмет исследования. В первой главе дан обзор современных методов обработки больших медицинских данных. Во второй главе обоснован и реализован метод NAUS. В третьей главе разработаны и оценены пять прогностических моделей. В четвертой главе описана архитектура программного пакета и интеграция в него обученных моделей.
		1) отражает;	
		2) частично отражает;	
		3) не отражает.	
		4.3. Цель и задачи соответствуют теме диссертации:	Цель и задачи исследования сформулированы корректно и полностью соответствуют теме диссертации. Все поставленные задачи логично подчинены достижению основной цели - повышению точности прогнозирования клинических исходов почечной трансплантации за счет специального метода балансировки данных и моделей машинного обучения.
		1) соответствуют;	
		2) частично соответствуют;	
		3) не соответствуют.	

		4.4 Все разделы и положения диссертации логически взаимосвязаны: 1) полностью взаимосвязаны; 2) взаимосвязь частичная; 3) взаимосвязь отсутствует.	Все разделы и положения диссертации находятся в логической взаимосвязи. Теоретический обзор в первой главе формирует основу для постановки задачи, во второй главе предложен метод подготовки данных, в третьей главе реализованы прогностические модели, а в четвертой главе показано, как эти модели объединены в программном пакете. Такая структура обеспечивает внутреннее единство работы и последовательность изложения.
		4.5 Предложенные автором новые решения (принципы, методы) аргументированы и оценены по сравнению с известными решениями: 1) критический анализ есть; 2) анализ частичный; 3) анализ представляет собой не собственные мнения, а цитаты других авторов; 4) анализ отсутствует.	Предложенные автором решения аргументированы и сопоставлены с известными подходами. Метод NAUS сравнивается не только с существующими вариантами undersampling, но и с oversampling-подходами. Эффективность оценивается по совокупности метрик Accuracy, AUC, Precision, Recall, F1-score, AUCPR и Brier score, что позволяет всесторонне оценить качество полученных результатов и подтвердить преимущества предложенного решения.
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%).	Основные научные результаты диссертации являются новыми. Ключевым результатом выступает разработка метода NAUS, предназначенного для балансировки шумных и несбалансированных медицинских данных. Метод ориентирован на снижение ошибок прогнозирования и может применяться к различным типам медицинских данных.
		5.2 Выводы диссертации являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%).	Выводы диссертации являются новыми, поскольку основаны на сочетании двух авторских результатов: метода NAUS и пяти моделей машинного обучения, интегрированных в программный пакет для прогнозирования клинического состояния после трансплантации. Существенным подтверждением новизны служит рост точности прогностических моделей в среднем на 16% при использовании предложенного метода балансировки.

		5.3 Технические, технологические, экономические или управленческие решения являются новыми и обоснованными:	Технологические решения, предложенные в диссертации, являются частично новыми и научно обоснованными. Речь идет о методе балансировки NAUS и программном пакете, в который интегрированы пять моделей машинного обучения для прогнозирования состояния пациента и жизнеспособности трансплантата на разных временных интервалах. Их обоснованность подтверждена результатами обработки медицинских данных, сравнительным анализом и оценкой по системе метрик.
		1) полностью новые;	
		2) частично новые (новыми являются 25-75%);	
		3) не новые (новыми являются менее 25%).	
6.	Обоснованность основных выводов	Все основные выводы основаны/не основаны на весомых с научной точки зрения доказательствах либо достаточно хорошо обоснованы (для qualitative research (куолитатив ресеч) и направлений подготовки по искусству и гуманитарным наукам).	Основные выводы работы основаны на результатах анализа данных, статистической обработки и вычислительных экспериментов. Достоверность обеспечивается сравнением с известными методами, применением контрольных медицинских наборов данных, кросс-валидацией и использованием количественных показателей качества моделей.
7.	Основные положения, выносимые на защиту	Необходимо ответить на следующие вопросы по каждому положению в отдельности:	7.1 Положения, выносимые на защиту, доказаны в достаточной степени. Эффективность метода NAUS подтверждена на контрольных наборах данных, а результативность пяти моделей машинного обучения с помощью кросс-валидации и метрик AUC и F1-score. 7.2 Указанные положения не являются тривиальными, поскольку автор не ограничивается применением готовых схем, а предлагает новый способ балансировки данных и самостоятельную реализацию пяти моделей для разных временных горизонтов. 7.3 Положения обладают научной новизной: новыми являются и метод NAUS, и комплекс прогностических моделей, и их объединение в программном пакете. 7.4 Уровень применения следует оценить как широкий, поскольку метод балансировки может быть использован для различных типов медицинских данных, а модели ориентированы на решение значимой задачи
		7.1 Доказано ли положение?	
		1) доказано;	
		2) скорее доказано;	
		3) скорее не доказано;	
		4) не доказано;	
		5) в текущей формулировке проверить доказанность положения невозможно.	
		7.2 Является ли тривиальным?	
		1) да;	
		2) нет;	
		3) в текущей формулировке проверить тривиальность положения невозможно.	
		7.3 Является ли новым?	

		1) да; 2) нет; 3) в текущей формулировке проверить новизну положения невозможно. 7.4 Уровень для применения: 1) узкий; 2) средний; 3) широкий; 4) в текущей формулировке проверить уровень применения положения невозможно. 7.5 Доказано ли в статье? 1) да; 2) нет; 3) в текущей формулировке проверить доказанность положения в статье невозможно.	посттрансплантационного прогноза. 7.5 Основные положения получили отражение в опубликованных научных работах автора.
8.	Принцип достоверности.	8.1 Выбор методологии - обоснован или методология достаточно подробно описана:	Выбор методологии является обоснованным и достаточно подробно описанным. В диссертации последовательно раскрыты особенности больших медицинских данных, проблема дисбаланса классов, логика разработки метода NAUS, подход к построению прогностических моделей и принципы их интеграции в программный пакет.
	Достоверность источников и предоставляемой информации	1) да;	
		2) нет.	
		8.2 Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий:	Результаты диссертации получены с использованием современных методов интеллектуального анализа данных, машинного обучения, Data Mining и статистической обработки. Применение вычислительных экспериментов и программных средств полностью соответствует характеру исследуемой задачи.
		1) да;	
		2) нет.	

		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):	Теоретические выводы и предложенные модели подтверждены экспериментально. Метод NAUS прошел проверку на медицинских контрольных наборах данных, а пять моделей машинного обучения оценены по различным временным горизонтам с применением кросс-валидации и показателей AUC и F1-score.
		1) да;	
		2) нет.	
		8.4 Важные утверждения подтверждены/частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу.	Ключевые утверждения работы подтверждены ссылками на актуальные и достоверные научные источники. В литературном обзоре рассмотрены современные подходы к обработке больших медицинских данных, методы работы с дисбалансом классов, а также особенности прогнозирования в трансплантологии.
		8.5 Использованные источники литературы достаточны/не достаточны для литературного обзора.	Использованный список литературы является достаточным для выполнения качественного литературного обзора и формирования надежной теоретико-методологической базы исследования.
9	Принцип практической ценности	9.1 Диссертация имеет теоретическое значение:	Теоретическая значимость диссертации состоит в развитии подходов к обработке несбалансированных больших медицинских данных и в разработке алгоритма NAUS, объединяющего фильтрацию шума и многостадийное undersampling. Работа также показывает значимость учета временного фактора при прогнозировании посттрансплантационных исходов.
		1) да;	
		2) нет.	
		9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике:	Практическая значимость диссертации высока. Метод NAUS повышает качество подготовки данных и способствует улучшению точности моделей. Пять интегрированных в программный пакет моделей могут использоваться для прогнозирования клинического состояния после трансплантации, включая оценку жизнеспособности
		1) да;	
		2) нет.	

			трансплантата на разных временных интервалах, и могут служить аналитическим инструментом поддержки принятия решений.
		9.3 Предложения для практики являются новыми:	Практические предложения, сформулированные в диссертации, являются частично новыми. Их новизна связана с применением специального метода балансировки медицинских данных и использованием программного пакета, в который встроены пять машинно-обучающих моделей для практического прогноза после трансплантации.
		1) полностью новые;	
		2) частично новые (новыми являются 25-75%);	
		3) не новые (новыми являются менее 25%).	
10.	Качество написания и оформления	Качество академического письма:	Диссертация написана качественно и оформлена на высоком академическом уровне. Текст логично выстроен, терминология употребляется корректно, а иллюстративный материал, таблицы и формулы способствуют восприятию результатов.
		1) высокое;	
		2) среднее;	
		3) ниже среднего;	
		4) низкое.	
11.	Замечания к диссертации	Представляется целесообразным более подробно описать механизмы возможной интеграции разработанного программного пакета с клиническими информационными системами. Также работу можно было бы усилить более развернутым описанием результатов внешней независимой валидации. Отмеченные замечания носят уточняющий характер и не снижают научной ценности диссертационной работы.	
12.	Научный уровень статей докторанта по теме исследования (в случае защиты диссертации в форме серии статей официальные рецензенты комментируют научный уровень каждой статьи докторанта по теме исследования)	Научный уровень публикаций докторанта по теме исследования следует оценить как высокий. По материалам диссертации опубликовано 7 научных работ, в том числе статьи в изданиях, рекомендованных КОКСНВО МНВО РК, публикации в журналах, индексируемых в Web of Science и/или Scopus, а также материалы международных конференций. Это соответствует установленным требованиям к диссертационным исследованиям.	

13.	Решение официального рецензента (согласно пункту 28 настоящего Типового положения)	Считаю, что Ерқос Айнұр заслуживает присуждения степени доктора философии (PhD) по образовательной программе «8D06102 – Компьютерная инженерия».

Официальный рецензент:
Директор НИИ «Искусственный интеллект» НАО «Евразийский национальный университет имени Л.Н. Гумилева», PhD, ассоциированный профессор



Барлыбаев Алибек Бактыбаевич

