

ОТЗЫВ ОФИЦИАЛЬНОГО РЕЦЕНЗЕНТА

на диссертационную работу Бакировой Гульназ Сайлауовны
на тему **«Разработка моделей и методов с применением федеративного машинного обучения»**,
представленную на соискание степени доктора философии (PhD)
по образовательной программе 8D06102 – «Компьютерная и программная 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 направленным развитию науки и/или государственным программам	<p>1.1 Соответствие направлению развитию 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) диссертация соответствует приоритетному направлению развития науки, утвержденному Высшей научно-технической комиссией при Правительстве Республики Казахстан (указать направление).</p>	<p>Диссертация соответствует целям следующих государственных программ:</p> <ul style="list-style-type: none"> • Национальный проект «Цифровой Казахстан», утвержденный постановлением Правительства Республики Казахстан от 12 декабря 2017 года № 827. • Концепция развития науки в Республике Казахстан на 2022–2026 годы, утвержденная Постановлением Правительства Республики Казахстан от 19 октября 2021 года № 736. • Программа «Цифровое правительство» (в рамках «Цифрового Казахстана») исследование связано с применением технологий искусственного интеллекта и может быть использовано при разработке интеллектуальных систем обработки и анализа данных. • Стратегический план развития искусственного интеллекта до 2030 года, инициированный в рамках реализации поручений Президента на V заседании Национального курултая от 15 января 2026 года.
2	Важность для науки	Работа вносит/не вносит существенный вклад в науку, а ее важность <u>хорошо раскрыта/не раскрыта</u> .	Диссертационная работа вносит существенный вклад в развитие методов федеративного машинного обучения и обработки распределенных данных. Актуальность исследования обоснована, а предложенный подход к обучению моделей в условиях гетерогенных (по IID) данных отражает его практическую направленность.

		<p>Полученные результаты обладают прикладной значимостью и могут быть применены в системах, работающих с конфиденциальной и распределенной информацией, включая образовательные и корпоративные аналитические платформы.</p>
<p>3</p>	<p>Принцип самостоятельности</p>	<p>Уровень самостоятельности оценивается как высокий. Диссертационная работа выполнена автором самостоятельно на всех этапах проведения научного исследования.</p>
<p>4</p>	<p>Принцип внутреннего единства</p>	<p>4.1 Обоснование актуальности диссертации: 1) <u>обоснована</u>; 2) частично обоснована; 3) не обоснована.</p> <p>В диссертационной работе актуальность темы представлена всесторонне и убедительно. Автор показывает, что развитие методов федеративного машинного обучения для анализа распределенных и конфиденциальных данных является важной и востребованной задачей в условиях роста объемов данных и ужесточения требований к защите персональной информации.</p> <p>Отмечается, что существующие подходы зачастую ориентированы на централизованную обработку данных или предполагают их полную доступность, что ограничивает их применение в реальных условиях, где данные распределены и обладают гетерогенной природой (по ИД).</p> <p>Актуальность исследования дополнительно подтверждается необходимостью внедрения технологий конфиденциального машинного обучения и интеллектуального анализа данных в различных прикладных областях, включая образовательные и корпоративные системы, что соответствует современным направлениям развития цифровых технологий и задачам обеспечения безопасности данных.</p>

	<p>4.2 Содержание диссертации отражает тему диссертации:</p> <ol style="list-style-type: none"> 1) <u>отражает</u>; 2) частично отражает; 3) не отражает. 	<p>Структура диссертационной работы соответствует заявленной тематике и последовательно раскрывает содержание исследования. Все разделы выстроены в логической взаимосвязи, отражают поставленные цели и задачи и обеспечивают целостность изложения. несоответствий теме диссертации не выявлено.</p>
	<p>4.3. Цель и задачи соответствуют теме диссертации:</p> <ol style="list-style-type: none"> 1) <u>соответствуют</u>; 2) частично соответствуют; 3) не соответствуют. 	<p>Цель и задачи согласуются с тематикой диссертационного исследования. Каждая из задач взаимосвязана с общей целью и отражает отдельные этапы работы, включая анализ существующих решений, разработку предложенного подхода и его экспериментальную оценку, что обеспечивает научную обоснованность и практическую завершенность исследования.</p>
	<p>4.4 Все разделы и положения диссертации логически взаимосвязаны:</p> <ol style="list-style-type: none"> 1) <u>полностью взаимосвязаны</u>; 2) взаимосвязь частичная; 3) взаимосвязь отсутствует. 	<p>Все разделы и положения диссертационной работы логически взаимосвязаны и подчинены общей цели исследования. Структура работы выстроена в соответствии с внутренней логикой научного анализа и направлена на последовательное решение поставленных задач. Исследование организовано по принципу перехода от анализа существующих методов федеративного машинного обучения и обработки распределенных данных к разработке и экспериментальной оценке предложенного подхода, что обеспечивает целостность и последовательность изложения результатов.</p>
	<p>4.5 Предложенные автором новые решения (принципы, методы) аргументированы и оценены по сравнению с известными решениями:</p> <ol style="list-style-type: none"> 1) <u>критический анализ есть</u>; 2) анализ частичный; 	<p>В диссертационной работе автором проведен систематический анализ современных методов федеративного машинного обучения и ансамблевых моделей, включая подходы к агрегации параметров, обработке распределенных данных и построению локальных моделей.</p> <p>На основе сопоставления существующих решений и обобщения результатов исследований установлены их</p>

	<p>3) анализ представляет собой не собственные мнения, а цитаты других авторов; 4) анализ отсутствует.</p>	<p>достоинства и ограничения при работе с гетерогенными (non IID) данными. С учетом проведенного анализа автором предложен и теоретически обоснован модифицированный подход к федеративному обучению, основанный на агрегировании статистических характеристик локальных моделей в виде векторов важности признаков. Эффективность разработанного подхода подтверждена результатами экспериментальных исследований и сравнительного анализа с существующими методами, что свидетельствует о его практической применимости.</p>
<p>5 Принцип научной новизны</p>	<p>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%).</p>	<p>Научные результаты диссертационной работы обладают научной новизной и актуальностью. Научные результаты диссертационной работы являются частично новыми. Используемые алгоритмы федеративного обучения - FedAvg, FedOrt, FedProx - а также модель Random Forest Regression представляют собой усовершенствованные методы, широко описанные в научной литературе, при этом тематика применения федеративного обучения в образовательной среде также отражена в ряде существующих исследований. Вместе с тем предложенная автором адаптация указанных алгоритмов к недифференцируемой ансамблевой модели посредством агрегирования статистических представлений локальных моделей в виде векторов важностей признаков представляет самостоятельный прикладной вклад. Разработанная архитектура федеративного обучения, обеспечивающая конфиденциальную обработку данных без передачи исходной информации на сервер, апробирована на реальных распределённых данных, что подтверждает практическую состоятельность выполненной работы.</p>
	<p>5.2 Выводы диссертации являются новыми? 1) полностью новые;</p>	<p>Полученные в ходе исследования результаты представляют собой значимый научный вклад и вносят вклад в развитие методов федеративного машинного обучения и интеллектуальной обработки распределенных</p>

	<p>2) частично новые (новыми являются 25-75%);</p> <p>3) не новые (новыми являются менее 25%).</p>	<p>Данных. Автору удалось не только предложить собственный подход к агрегированию и анализу данных, но и применить современные методы машинного обучения к условиям работы с гетерогенными (по ИД) распределенными данными.</p> <p>Новизна работы реализуется как в разработке подхода к модификации федерального обучения на основе агрегирования статистических характеристик локальных моделей, так и в использовании экспериментальных данных и проведении сравнительного анализа с существующими методами. Представленные решения расширяют возможности интеллектуальных систем анализа данных и созданот основу для дальнейшего совершенствования распределенных методов федерального машинного обучения.</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>В диссертационной работе разработана веб-платформа для интегрированного сбора и обработки данных, предназначенная для анализа распределенных и конфиденциальных данных в условиях федерального машинного обучения. Система обеспечивает формирование уникального набора данных на основе информации, поступающей от пользователей, включая локально собираемые показатели и их последующую обработку без передачи исходных данных на сервер.</p> <p>Автор последовательно реализовал этапы предобработки данных, включая их структурирование и агрегирование, а также применение методов статистической обработки и анализа, что позволило сформировать устойчивую методологию работы с гетерогенными (по ИД) распределенными данными. Разработанная архитектура системы обеспечивает конфиденциальную обработку информации за счет передачи на сервер только агрегированных обновлений моделей, что исключает</p>

		<p>необходимость передачи исходных данных. Такой подход способствует сохранению конфиденциальности и повышает надежность обработки информации. Все сформулированные выводы основаны на объективных результатах экспериментальных исследований и подтверждаются полученными эмпирическими данными, что свидетельствует о достоверности и практической применимости предложенных решений.</p>
<p>6</p> <p>Обоснованность основных выводов</p>	<p>Все основные выводы основаны/не основаны на весомых с научной точки зрения доказательствах либо достаточно хорошо обоснованы</p>	<p>Основные выводы диссертационного исследования опираются на достаточную совокупность научно обоснованных данных. Автор демонстрирует, что полученные результаты основаны на экспериментальных исследованиях, выполненных с использованием современных методов машинного обучения и обработки данных.</p> <p>Примененные методы модификации федеративного обучения и агрегирования статистических представлений локальных моделей обеспечивают воспроизводимые результаты, что позволяет считать выводы исследования научно обоснованными. Представленные доказательства являются достаточными и методологически корректными для количественного исследования в области машинного обучения и распределенных интеллектуальных систем.</p>
<p>7</p> <p>Основные положения, выносимые на защиту</p>	<p>Необходимо ответить на следующие вопросы по каждому положению в отдельности: 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.</p>	<p>Основные положения, выносимые на защиту, являются обоснованными и обладают научной и практической значимостью; они описываются определенной степенью научной новизны и подтверждаются результатами, представленными в публикациях автора. Полученные результаты отражены в статьях и материалах международных конференций, включая издания, индексируемые в базах Scopus и Web of Science и относящиеся к квартилю Q3 (47 й перцентиль), а также в 4 статьях, включенных в перечень ККСНВО РК.</p>

	<p>7.2 Является ли тривиальным?</p> <p>1) да; 2) <u>нет</u>; 3) в текущей формулировке проверить тривиальность положения невозможно.</p> <p>7.3 Является ли новым?</p> <p>1) да; 2) нет; 3) в текущей формулировке проверить новизну положения невозможно.</p> <p>7.4 Уровень для применения:</p> <p>1) узкий; 2) средний; 3) <u>широкий</u>; 4) в текущей формулировке проверить уровень применения положения невозможно.</p> <p>7.5 Доказано ли в статье?</p> <p>1) <u>да</u>; 2) нет; 3) в текущей формулировке проверить доказанность положения в статье невозможно.</p>	<p>Научная и практическая значимость результатов дополнительно подтверждена наличием авторского свидетельства Республики Казахстан № 62530 на программный продукт, опубликована 30.09.2025.</p>
<p>8 Принцип достоверности. Достоверность источников и предоставляемой информации</p>	<p>8.1 Выбор методологии обоснован или методология достаточно подробно описана:</p> <p>1) <u>да</u>; 2) нет.</p>	<p>Методология исследования обоснована и подробно описана: выбор архитектуры, методов агрегации и подходов к обучению обусловлен особенностями федеративного машинного обучения и задачами диссертации. Все этапы от разработки и модификации модели до экспериментальной оценки изложены последовательно и сопровождаются аргументацией в пользу реализованных решений, что обеспечивает воспроизводимость и научную обоснованность полученных результатов.</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) да; 2) нет.</p>	<p>Результаты диссертационного исследования получены с использованием современных методов научного анализа, включая методы машинного обучения и подходы федеративного обучения. При разработке моделей и анализе данных применялись надежные статистические методы оценки качества, а также современные инструментальные средства для обработки и анализа распределенных данных. Все этапы исследования реализованы с использованием современных компьютерных технологий, что обеспечило высокую точность, воспроизводимость и практическую значимость полученных результатов.</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) да; 2) нет.</p>	<p>Теоретические выводы, разработанные модели и выявленные взаимосвязи в рамках диссертационного исследования доказаны и подтверждены результатами экспериментальных исследований. Предложенные решения, включая подход к модификации федеративного машинного обучения для модели Random Forest Regression и архитектуру агрегации локальных моделей, прошли практическую проверку на соответствующих наборах распределенных данных.</p> <p>Результаты проверены с использованием общепринятых метрик качества и сопоставлены с существующими подходами, что подтверждает достоверность и научную обоснованность полученных результатов.</p>
	<p>8.4 Важные утверждения подтверждены/частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу.</p> <p>8.5 Исползованные источники литературы достаточны/недостаточны для литературного обзора.</p>	<p>Важные утверждения подтверждены ссылками на актуальную и достоверную научную литературу</p> <p>Исползованные источники литературы являются достаточными и охватывают широкий спектр современных научных работ в области федеративного машинного обучения, ансамблевых методов, анализа распределенных</p>

		<p>Данных и защиты конфиденциальности. В литературном обзоре представлены как классические подходы, так и актуальные исследования, включая публикации в высокорейтинговых научных журналах и материалах ведущих конференций. Это обеспечивает теоретическую обоснованность и актуальность проведенного диссертационного исследования.</p>
<p>9 Принципы практической ценности</p>	<p>9.1 Диссертация имеет теоретическое значение: 1) да; 2) нет.</p>	<p>Диссертация имеет теоретическое значение, поскольку в ней разработаны и обоснованы новые подходы к развитию федеративного машинного обучения для недифференцируемых моделей на примере Random Forest Regression. Предложена замена классической агрегации параметров на агрегирование статистических представлений локальных моделей, что расширяет теоретические представления о построении устойчивых моделей в условиях гетерогенных (non IID) распределенных данных.</p> <p>Сформулированные положения дополняют современные подходы к организации федеративного обучения и закладывают основу для дальнейших исследований в области распределенной обработки данных, конфиденциального машинного обучения и разработки устойчивых алгоритмов для реальных прикладных систем.</p>

	<p>9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике:</p> <p>1) да;</p> <p>2) нет.</p>	<p>Диссертация имеет практическое значение, поскольку в ее рамках разработан подход к расширению методов федеративного машинного обучения, адаптированный для модели Random Forest Regression. Предложено агрегирование статистических представлений локальных моделей (векторов важности признаков), что обеспечивает устойчивость при гетерогенных (non IID) данных.</p> <p>Разработанная архитектура позволяет формировать глобальную модель без передачи исходных данных клиентов на сервер, обеспечивая конфиденциальность. Полученные результаты применимы в распределенных системах и подтверждают практическую значимость и масштабируемость предложенных решений.</p>
<p>9.3 Предложения для практики 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>Предложенные в диссертационной работе решения для практического применения, являются полностью новыми, так как, основаны на авторских разработках в области федеративного машинного обучения. Разработан подход к модификации федеративного обучения для модели Random Forest Regression, основанный на агрегировании статистических представлений локальных моделей, что обеспечивает устойчивость при гетерогенных non IID данных. Разработаны методы федеративного обучения для анализа распределенных конфиденциальных данных, обеспечивающие локальную обработку информации и сохранение конфиденциальности. Предложена архитектура, позволяющая формировать глобальную модель без передачи исходных данных на сервер. Таким образом, решения имеют практическую значимость и эффективны для применения в распределенных системах.</p>	

10	<p>Качество написания и оформления</p>	<p>Качество академического письма:</p> <p>1) <u>высокое</u>; 2) <u>среднее</u>; 3) <u>ниже среднего</u>; 4) <u>низкое</u>.</p> <p>Качество академического изложения в диссертационной работе можно оценить как высокое, полностью соответствующее требованиям, предъявляемым к научным исследованиям. Материал представлен логично, последовательно и с точностью в формулировках. Применяется научный стиль речи с корректным использованием терминологии в областях лингвистики, компьютерных наук и искусственного интеллекта. Все утверждения аргументированы, выводы обоснованы, а структура текста способствует четкому и понятному раскрытию темы. В результате высокий уровень академического письма обеспечивает убедительность и хорошее восприятие содержания диссертации.</p>
11	<p>Замечания к диссертации</p>	<p>1. В Приложении Г, таблицы Г.3-Г.7) содержат – 50 наблюдений. Известно, что этих наблюдений было больше количество. С учетом этого, требуется явно указать конечное количество наблюдений. В противном случае, формируется мнение что ML-модели были обучены на таком малом количестве данных.</p> <p>2. Таблицы Г.8 (Ridge) и Г.9 (CatBoost) идентичны. Реальные значения, прогнозы и остатки в обеих таблицах полностью совпадают. Это требует исправления и/или объяснения.</p> <p>3. В тексте диссертации выявлены систематические орфографические, грамматические и стилистические ошибки. В частности, слова «объём», «объект», «объединение» последовательно написаны без твёрдого знака на протяжении всей работы. Встречаются опечатки: «перегрузке» вместо «перегрузке», «нагрузку» вместо «нагрузку», «фактор» вместо «фактор», «выполняет» вместо «выполняет», «континциальность» вместо «континциальность» и ряд других. Отмечены случаи несогласования падежей и рода в предложениях, использование разговорных конструкций, неуместных в академическом тексте («шпос объём выборки был ограничен»), а также аббревиатура «SMPS» вместо «SMPC», не соответствующая списку сокращений.</p>
12	<p>Научный уровень статей докторанта по теме исследования (в случае защиты диссертации в форме серии статей официальные рецензенты комментируют</p>	<p>Научный уровень публикаций соискателя полностью соответствует предъявляемым требованиям к диссертационным работам уровня PhD</p>

	научный уровень каждой статьи докторанта по теме исследования)	
13	Решение официального рецензента (согласно пункту 28 настоящего Типового положения)	Ходатайствовать перед Комитетом для присуждения докторанту степени доктора философии (PhD)

Заключение

Представленная диссертация на соискание степени доктора философии (PhD) **Бакировой Гульназ Сайлауовны** на тему «Разработка моделей и методов с применением федеративного машинного обучения», является законченной самостоятельной научно-исследовательской работой, которая отвечает требованиям «Правил присуждения степеней», а ее автор заслуживает ходатайствовать перед Комитетом о присуждении степени доктора философии (PhD).

Официальный рецензент:

к.т.н., профессор
Школы программной инженерии
Astana IT University, Астана





Амиргалиев Бейбүт Едилханович