ОТЗЫВ

официального рецензента на диссертационную работу
Шайзат Медета Жанболатұлы на тему «Разработка методов и моделей машинного обучения для киберкриминалистического анализа Bitcoin-транзакций», предоставленную на соискание степени доктора философии (PhD)
по образовательной программе «8D06301 – Системы информационной безопасности»

№ 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).	Диссертационное исследование в рамках грантового финансирования молодых ученых «Жас ғалым» на 2025-2027 годы по проекту AP25793719 «Разработка системы анализа нелегальных криптовалютных транзакций». Тематика проекта относится к приоритетному направлению «Национальная безопасность и оборона, биологическая безопасность» и специализированному направлению «Обеспечение информационной безопасности».
2.	Важность для науки	Работа вносит/не вносит существенный вклад в науку, а ее важность хорошо раскрыта/не раскрыта.	Научная значимость работы определяется решением актуальной для блокчейн-аналитики проблемы: совмещением высокого качества классификации с содержательной объяснимостью полученного результата. Автор не ограничивается выбором алгоритма машинного обучения, а последовательно рассматривает состав данных, временную структуру транзакционного графа,

			воспроизводимость признаков и возможность их криминалистической интерпретации. Существенный вклад в развитие данного научного направления вносят система семантически открытых графовых и временных дескрипторов, модель выявления нелегальных Bitcoin-транзакций с использованием разработанных дескрипторов и метод отбора признаков для наборов данных с анонимизированными признаками. Полученные результаты расширяют методическую базу применения машинного обучения в цифровой криминалистике и финансовом мониторинге.
3.	Принцип самостоятельности	Уровень самостоятельности: 1) высокий; 2) средний; 3) низкий; 4) самостоятельности нет.	Уровень самостоятельности соискателя оценивается как высокий. Автором выполнен полный исследовательский цикл: от постановки проблемы, восстановления идентификаторов и подготовки данных до формализации признаков, обучения моделей, анализа ошибок и программной реализации. Обоснование принятых решений и критическая оценка ограничений подтверждают самостоятельный характер работы.
4.	Принцип внутреннего единства	4.1 Обоснование актуальности диссертации: 1) обоснована; 2) частично обоснована; 3) не обоснована.	Актуальность исследования обоснована сочетанием практических и научных факторов. Публичность Bitcoin позволяет прослеживать движение средств, но псевдонимность участников, неполная разметка и сложность графового контекста затрудняют интерпретацию автоматизированных выводов. Поэтому востребованы модели, связывающие прогноз с наблюдаемыми характеристиками транзакции и пригодные для финансового мониторинга.
		4.2 Содержание диссертации отражает тему диссертации: 1) отражает; 2) частично отражает; 3) не отражает.	Содержание полностью соответствует теме. Первый раздел раскрывает основы анализа Bitcoin-транзакций, второй посвящен отбору анонимизированных признаков и сравнению моделей, третий содержит разработку открытых дескрипторов и модели CatBoost, четвертый

			представляет гибридную программную систему. Все заявленные в названии методы, модели и прикладные средства последовательно рассмотрены.
		4.3. Цель и задачи соответствуют теме диссертации: 1) соответствуют; 2) частично соответствуют; 3) не соответствуют.	Цель исследования непосредственно связана с темой и направлена на разработку интерпретируемых признаков и моделей машинного обучения. Восемь задач охватывают анализ подходов и данных, отбор и формирование признаков, сравнение моделей, экспериментальную проверку и создание программной системы. Представленные результаты подтверждают решение задач и достижение цели.
		4.4 Все разделы и положения диссертации логически взаимосвязаны: 1) полностью взаимосвязаны; 2) взаимосвязь частичная; 3) взаимосвязь отсутствует.	Разделы и положения логически взаимосвязаны. Ограничения анонимизированных признаков, выявленные во втором разделе, обосновывают переход к воспроизводимым ончейн-дескрипторам в третьем. Полученная модель далее включается в гибридную систему. Тем самым каждый последующий этап опирается на результаты предыдущего.
		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 решения сопоставлены с известными подходами в единых экспериментальных условиях. Генетический отбор сравнивается с полным набором, SelectKBest, опубликованными результатами и GNN. Интерпретируемое представление проверено на восьми ансамблевых, графовых и гибридных моделях. Анализ временного разделения, частичной разметки, информационной утечки, абляций и ошибок свидетельствует о полноценном критическом анализе.
5.	Принцип научной новизны	5.1 Научные результаты и положения являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%);	Научные результаты и положения обладают новизной. Предложены: система из 61 интерпретируемого дескриптора, модель CatBoost на 87 открытых признаках без анонимизированных переменных Elliptic и оберточный генетический отбор признаков с оценкой Random Forest.

		3) не новые (новыми являются менее 25%).	Обоснованность и корректность научных результатов, положений подтверждаются их апробацией на международных научных конференциях и публикацией результатов в рецензируемых изданиях, включая журналы, индексируемые в базе Scopus и рекомендованные КОКСНВО МНВО РК.
		5.2 Выводы диссертации являются новыми?	Выводы диссертации являются новыми и имеют самостоятельный научный характер. Экспериментально показано, что качество классификации зависит не только от алгоритма, но и от семантики признаков, построения графа, временного протокола и правил работы с частичной разметкой. Дополнительная информативность разработанных дескрипторов подтверждена для всех восьми моделей.
		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:	Новым техническим решением является модульная система, объединяющая расчет признаков и классификацию, проверку адресов, поведенческие индикаторы и локальный графовый анализ. Веб-приложение с REST API и базой данных сохраняет результаты модулей раздельно, формирует аналитический кейс и поддерживает проверяемую интерпретацию без подмены экспертной оценки.
		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 хорошо обоснованы. Эксперименты проведены на реальных данных Elliptic и Elliptic++, восстановленных txid и открытых ончейн-характеристиках. Соблюдены временные границы, исключено использование тестовых срезов при отборе признаков, применены единые разбиения для сравнения моделей. Достоверность дополнительно подтверждают повторные разбиения графов, абляционный анализ, матрицы ошибок, SHAP-интерпретация и оценка вычислительной эффективности.

7.	Основные положения, выносимые на защиту	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. 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;	**Положение № 1.** Система криминалистически интерпретируемых признаков Bitcoin-транзакций, формализующая криминалистические гипотезы, повышает качество выявления нелегальных транзакций. Положение доказано. Выполнено сравнение базового набора из 26 признаков и расширенного набора из 87 признаков. Добавление 61 разработанного дескриптора обеспечило повышение PR-AUC и F1-score у всех восьми проверенных моделей. Результаты абляционного анализа дополнительно подтвердили вклад признаков локального отклонения от окружения, последующего расходования средств и предшествующей транзакционной истории в качество классификации. Положение не является тривиальным, поскольку преобразование криминалистических гипотез в систему формально определенных и вычисляемых количественных признаков требует объединения методов анализа транзакционных графов, временной динамики и машинного обучения. Новизна положения состоит в формировании системы признаков с открытой семантикой, позволяющей содержательно интерпретировать причины отнесения Bitcoin-транзакции к определенному классу. Уровень применения положения оценивается как широкий, поскольку разработанные признаки могут использоваться в системах финансового мониторинга, правоохранительными органами, криптовалютными биржами и организациями, осуществляющими анализ операций с цифровыми активами. Доказано в статье: M. Shaizat, Sh. Mussiraliyeva, I. Tereikovskyi, «Forensically Interpretable Graph Descriptors for Improved Illicit Bitcoin Transaction Detection», International Journal of Computer Network and Information Security, Vol. 18, No. 3, pp. 63-84, 2026.

		3) в текущей формулировке проверить доказанность положения в статье невозможно.	**Положение № 2.** Модель выявления нелегальных Bitcoin-транзакций на основе алгоритма CatBoost обучена на сформированном наборе данных с семантически открытыми транзакционными, структурно-графовыми и криминалистически интерпретируемыми признаками. Положение доказано. Модель CatBoost показала лучшие значения среди рассмотренных конфигураций: PR-AUC составил 90,27%, F1-score - 83,90%, Recall - 81,11%. Полученные показатели подтверждают работоспособность предложенной модели и целесообразность совместного использования различных групп признаков. Положение не является тривиальным. Реализация потребовала восстановления публичных идентификаторов транзакций из анонимизированного набора Elliptic, формирования нового признакового пространства из 87 характеристик и настройки модели с учетом дисбаланса классов и пропущенных значений. Новизна положения заключается в разработке модели, не зависящей от анонимизированных признаков Elliptic и допускающей расчет входных характеристик по открытым данным блокчейна Bitcoin. Уровень применения положения оценивается как широкий, поскольку модель может быть интегрирована в программные системы мониторинга криптовалютных операций и использоваться для первичного отбора транзакций, требующих дальнейшей экспертной проверки. Доказано в статье: M. Shaizat, Sh. Mussiraliyeva, I. Tereikovskyi, «Forensically Interpretable Graph Descriptors for Improved Illicit Bitcoin Transaction Detection», International Journal of Computer Network and Information Security, Vol. 18, No. 3, pp. 63-84, 2026.

			Положение № 3. Метод отбора признаков для блокчейн-наборов данных с анонимизированными признаками основан на генетическом поиске подмножеств и их оберточной оценке моделью Random Forest. Положение доказано. Эксперименты проводились на наборах данных Elliptic и Elliptic++. Для Elliptic число признаков было сокращено со 166 до 53, при этом F1-score повысился с 82,90% до 84,26%. Для Elliptic++ число признаков сократилось со 183 до 85, а F1-score повысился с 82,50% до 84,07%. При одинаковом количестве признаков генетический отбор также превзошел фильтрационный метод SelectKBest. Положение не является тривиальным. Генетический поиск требует многократного обучения модели для оценки различных сочетаний признаков и должен выполняться с соблюдением временного разделения данных, исключающего утечку информации из тестовых срезов. Новизна положения заключается в применении оберточного генетического отбора к блокчейн-наборам с анонимизированными признаками, содержательный экспертный анализ которых невозможен. Уровень применения оценивается как широкий. Метод полезен для исследователей и разработчиков моделей машинного обучения, работающих с высокоразмерными анонимизированными наборами данных. Положение опубликовано в статье: Shaizat M., Mussiraliyeva S. «Enhanced Identification of Illicit Bitcoin Transactions through Genetic Algorithm-Based Feature Selection» // Eastern-European Journal of Enterprise Technologies. 2025. Vol. 4, No. 9 (136). P. 34-42.
8.	Принцип достоверности.	8.1 Выбор методологии - обоснован или методология достаточно подробно описана:	Методология соответствует цели исследования. Сравнительный анализ, генетическая оптимизация, робастная статистика, графовая авторегрессия,

	Достоверность источников и предоставляемой информации	1) да;	ансамблевые и графовые модели применены в соответствии с решаемыми задачами. Правила формирования выборок, метрики, параметры и обработка пропусков описаны достаточно подробно.
		2) нет.	
		8.2 Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий:	Работа выполнена с применением современных методов и технологий: ансамблевого машинного обучения, GNN, генетического поиска, SHAP, графовых метрик, API блокчейн-сервисов и программной обработки транзакционных данных. Система реализована на Python с использованием Django, Django REST Framework, PostgreSQL, NetworkX и профильных библиотек машинного обучения.
		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 положения и модели подтверждены вычислительными экспериментами. Метод отбора проверен на Elliptic и Elliptic++, а система дескрипторов исследована на восьми классификаторах, графовых и гибридных моделях. Снижение метрик при абляции отдельных групп и результаты SHAP согласуются с заявленной интерпретацией признаков.
		1) да;	
		2) нет.	
		8.4 Важные утверждения подтверждены/частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу.	Ключевые утверждения подтверждены ссылками на актуальные научные публикации по цифровой криминалистике, блокчейн-графам и машинному обучению, также нормативными актами, документами FATF и FinCEN, официальными материалами расследований. Автор учитывает методологические различия и ограничения цитируемых исследований.
		8.5 Использованные источники литературы достаточны/не достаточны для литературного обзора.	Список из 104 источников достаточен для литературного обзора. Он включает фундаментальные и современные работы по Bitcoin, графовому анализу, Elliptic и Elliptic++, выявлению нелегальных операций, отбору признаков,

			интерпретации моделей и международному регулированию виртуальных активов.
9	Принцип практической ценности	9.1 Диссертация имеет теоретическое значение: 1) да; 2) нет.	Теоретическое значение состоит в развитии интерпретируемого анализа Bitcoin-транзакций. Разработанные дескрипторы формализуют локальное отклонение, динамику последующего расходования и связь текущих параметров с восходящей историей. Работа показывает способ перевода криминалистических гипотез в воспроизводимые количественные признаки.
		9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике: 1) да; 2) нет.	Практическая ценность подтверждается созданием инструментария, который формирует признаковое и графовое представление транзакции, выполняет классификацию, адресную проверку и анализ поведения.
		9.3 Предложения для практики являются новыми: 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%).	Практические предложения являются полностью новыми в рамках разработанной системы. Использование 87 семантически открытых и интерпретируемых признаков вместо анонимизированного признакового пространства позволяет не только классифицировать Bitcoin-транзакции, но и дополнять результаты проверкой связанных адресов, анализом их активности и локального графового окружения.
10.	Качество написания и оформления	Качество академического письма: 1) высокое; 2) среднее; 3) ниже среднего; 4) низкое.	Качество академического письма высокое. Материал изложен последовательно, терминология используется единообразно, математические обозначения пояснены, выводы согласованы с задачами и экспериментами. Таблицы, рисунки и схемы дополняют текст. Представленная диссертационная работа полностью соответствует установленным требованиям к оформлению и содержанию исследований на соискание степени доктора философии (PhD).

11.	Замечания к диссертации	По диссертационной работе можно высказать следующие замечания и рекомендации: *1.* В описании программной системы одновременно используются русскоязычные названия модулей и англоязычные обозначения элементов пользовательского интерфейса, что создает небольшую терминологическую неоднородность *2.* В работе присутствуют грамматические и стилистические ошибки. Перечисленные замечания носят рекомендательный характер, не затрагивают достоверность основных результатов и не снижают высокой научной и практической оценки диссертации.
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 соответствует положениям диссертации. Опубликовано пять работ: две статьи в журналах Scopus, две статьи в изданиях, рекомендованных КОКСНВО, и одна работа в материалах международной конференции. Статья по генетическому отбору раскрывает третье положение, а публикация по интерпретируемым дескрипторам отражает первое и второе положения. Процентили Scopus-журналов 57 и 72 подтверждают достойный уровень апробации.
13.	Решение официального рецензента (согласно пункту 28 настоящего Типового положения)	Диссертационная работа Шайзат Медета Жанболатұлы «Разработка методов и моделей машинного обучения для киберкриминалистического анализа Bitcoin-транзакций» является завершенным самостоятельным исследованием высокого научно-методического уровня. Работа соответствует требованиям к диссертациям на соискание степени доктора философии (PhD), а автор заслуживает присуждения степени доктора философии (PhD) по образовательной программе «8D06301 - Системы информационной безопасности».

Официальный рецензент,
доктор философии (PhD),
главный ученый секретарь,
ассоциированный профессор
Института информационных и вычислительных технологий
Комитета науки Министерства науки и высшего образования
Республики Казахстан

Усатова О.А.