DOI: 10.47026/2499-9636-2024-4-1-14

УДК 336.717:343.72+37:004
ББК У262.5р.с51+Х408.121.12

О.Г. АРКАДЬЕВА

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОТСЛЕЖИВАНИЯ СОЦИАЛЬНОЙ ИНЖЕНЕРИИ В БАНКОВСКИХ ТРАНЗАКЦИЯХ

Ключевые слова: *дисбаланс данных, асимметрия, передискретизация, дрейф концепций, ранняя остановка, модели машинного обучения.*

Актуальность исследования обусловлена тем, что в эпоху цифровизации и повсеместного внедрения технологий онлайн-платежи и иные дистанционные банковские транзакции становятся все более популярными, что приводит к росту случаев мошенничества с использованием социальной инженерии.

Цели исследования *– оценка возможностей методов машинного обучения для противодействия социальной инженерии, а также выявление ограничений использования этих методов в банковских антифрод-системах.*

Материалы и методы*. Для проверки гипотезы о недостаточной эффективности методов машинного обучения для распознавания атак с использованием социальной инженерии использовался набор данных Bank Account Fraud Dataset Suite (NeurIPS 2022). Подготовка данных для машинного обучения была реализована с помощью алгоритма нелинейного уменьшения размерности UMAP и метода синтетической передискретизации SMOTE.*

Результаты исследования*. Кластеризация данных не позволила достаточно достоверно разделить мошеннические и немошеннические операции в зависимости от каких-либо характеристик клиента или самой транзакции. При этом точность модели на основе деревьев решений наивысшая по сравнению с градиентным бустингом и логистической регрессией, однако эффективность и результативность практического применения алгоритмов неизбежно снижается из-за дрейфа концепции, вызывающего увеличение вероятности ошибочной классификации. В случае дрейфа концепции модель становится переобученной и плохо работает на тестовых данных. Для регуляризации может использоваться ранняя остановка, когда процесс обучения модели должен остановиться в точке, где потери и ошибки в проверочных данных достигают минимального значения.*

Выводы*. Поскольку банки обязаны противодействовать переводам, которые происходят без добровольного согласия клиента, в частности под воздействием злоумышленников, антифрод-системы банков-отправителей и банков-получателей платежей должны регулярно переобучаться во избежание распространения мошеннических случаев и получения банками убытков от необходимости осуществления компенсационных выплат. Практическая значимость заключается в возможности использования результатов для отслеживания и противодействия социальной инженерии в банковских транзакциях в целях дальнейшего совершенствования регуляторных усилий и действий самих коммерческих банков.*

Введение. В эпоху цифровизации и повсеместного внедрения технологий онлайн-платежи и иные банковские транзакции в дистанционном формате становятся все более популярными. Y.-F. Lin et al. подчеркивают, что из-за пандемии COVID-19 банковские транзакции все чаще совершаются без личного

присутствия клиентов. Однако эта тенденция также предоставила злоумышленникам возможность получить информацию о персональных данных клиентов различными незаконными способами, что привело к росту случаев мошенничества с использованием социальной инженерии [15]. Вследствие активизации социальной инженерии финансовые организации несут как репутационные, так и прямые финансовые потери, а население лишается сбережений и утрачивает доверие к финансовым институтам [10]. Существует необходимость в развитии методов анализа и обнаружения мошеннических транзакций, обусловленных социальной инженерией, и отделения их от законных, дополнения их мерами государственного регулирования [1]. В основном на практике для обнаружения мошенничества используются подходы контролируемого обучения, основывающиеся на сочетании признаков мошенничества, которые можно выявить с помощью анализа прошлых транзакций. Однако эти подходы неэффективны, когда происходят изменения в поведении клиентов, вызванных психологическим воздействием на них. Более того, сложно выявить аномальные транзакции из-за значительного дисбаланса данных о мошеннических и немошеннических операциях.

Целью данного исследования является оценка возможностей методов машинного обучения для противодействия социальной инженерии, а также выявление ограничений использования этих методов в банковских антифрод-системах.

Материалы и методы. Для проверки гипотезы о недостаточной эффективности методов машинного обучения в случае использования социальной инженерии использовался набор данных Bank Account Fraud Dataset Suite (NeurIPS 2022). В целях сравнения методов машинного обучения категориальные признаки были преобразованы в числовые значения, было выполнено логарифмическое преобразование целевого признака ввиду значительной степени смещения, в отношении набора данных был реализован алгоритм нелинейного уменьшения размерности UMAP и метод синтетической передискретизации SMOTE.

Результаты исследования. Проблема мошенничества при совершении банковских транзакций активно изучается в научной среде. V. Rutskiy делает выводы о необходимости создания инструментов банковской аналитики для рынка в целом и разработки систем сбора данных для работы алгоритмов искусственного интеллекта, совершенствования нормативной базы данной сферы банковского дела и повышения осведомленности клиентов банка о возможностях технологии искусственного интеллекта [18]. P. Vanini et al. определяют три основные модели для борьбы с мошенническими банковскими транзакциями: обнаружение мошенничества на основе машинного обучения, экономическая оптимизация результатов машинного обучения и модель риска для прогнозирования риска мошенничества с учетом мер противодействия [22]. A. Vashistha et al. подчёркивают, что методы контролируемого машинного обучения имеют ограничения, такие как высокий уровень ложных срабатываний, низкая масштабируемость и уязвимость к состязательным атакам, и предлагают алгоритм гиперансамблевого машинного обучения [23]. Исследователями V.L.N. Gorle et al. представлена полуконтролируемая модель борьбы

с мошенничеством на основе выбросов, позволяющая идентифицировать заявителя на получение кредита как подлинного или мошеннического должника [12]. A.I. Sergadeeva et al. представляют финансовые транзакции в виде графов и анализируют их с помощью графовой нейронной сети с целью обнаружения транзакций, типичных для схем мошенничества [20]. GR J. et al. для обнаружения финансового мошенничества предлагают использование настроенной двунаправленной LSTM-модели с уровнем внимания (A-BiLSTM) [11]. P. Boulieris et al. утверждают, что онлайн-действия, действительно, подчиняются правилам, аналогичным правилам естественного языка и, следовательно, к ним можно успешно подойти с помощью методов обработки естественного языка [9].

A.M. Salam et al. подчеркивают, что существует значительный дисбаланс в операциях по кредитным картам во всех банках: небольшой процент мошеннических транзакций превышает большинство действительных, поэтому набор данных должен быть сбалансирован, и акцент делается ими на сравнительном анализе нескольких индивидуальных и гибридных методов повторной выборки [19]. G. Zioviris et al. представили попытку создать гибридную систему, т.е. последовательную схему объединения двух моделей глубокого обучения и эффективного определения финансовых операций с использованием метода передискретизации. Передискретизация применяется для обработки сильно несбалансированных наборов данных, что является «естественным» сценарием из-за ограниченного количества мошенничеств по сравнению с объемами транзакций [25].

A. Vashistha et al. изучили применение глубокой нейронной сети, модели опорных векторов, многослойного персептрона, K-ближайших соседей, случайного леса, алгоритмов XGBoost, LGBM и деревьев решений, из которых случайный лес, дерево решений, XGBoost и LGBM показали 100% точность за счет балансировки набора данных [24]. M. Karbasiyan et al. отобрали алгоритмы XGBoost и LightGBM в соответствии с высоким ROC в полученных моделях [14]. P. Hajek et al. отмечают, что полуконтролируемая ансамблевая модель, объединяющая несколько неконтролируемых алгоритмов обнаружения выбросов и классификатор XGBoost, достигает наилучших результатов, в то время как наибольшая экономия средств может быть достигнута за счет комбинирования методов случайной выборки и XGBoost [13]. S.R. Bin et al. предлагают гибридное решение, использующее нейронную сеть в рамках федеративного обучения [8].

Из соображений безопасности и конфиденциальности данных банки не могут разглашать данные о транзакциях и не делают этого даже в виде обезличенных наборов данных. Эта проблема существенно затрудняет разработку и развитие моделей обнаружения мошенничества. T.K. Dang et al. предлагают такой выход: банки по-прежнему обучают локальные антифрод-модели, используя свои собственные базы данных. Впоследствии создается общая глобальная модель путем агрегирования локально рассчитанных обновлений модели обнаружения мошенничества, что позволяет банкам коллективно извлекать выгоду из сотрудничества без обмена наборами данных и защищает конфиденциальную информацию клиентов [13].

Разработки российских ученых сконцентрированы на совершенствовании алгоритмов машинного обучения и совершенствования правовых, организационных, психологических, информационных и других компонентов обеспечения безопасности банковских транзакций [2, 3, 5, 7].

Мегарегулятор, представляя статистические данные о мошеннических действиях при совершении банковских транзакций, оперирует термином «ОБС – операции без согласия» (табл. 1).

Таблица 1

Инциденты информационной безопасности при переводе денежных средств физическими лицами в первом квартале 2024 г.

Способ перевода	Количество ОБС, ед.	Объем ОБС, тыс. руб.	Удельный вес возмещенных (возвращенных) средств, %
Карты	237 207	1 918 855,52	8,4
Счета (дистанционное банковское обслуживание, переводы)	13 717	901 694,62	16,2
СБП	41 458	1 131 339,16	1,5
Электронные кошельки	1 651	39 884,93	0
Без открытия счета	81	5 963,16	0

Примечание. Табл. 1 составлена автором по данным [4].

Удельный вес возмещенных (возвращенных) средств в I квартале 2024 г. составил 7,7% против 8,7% в среднем за 2023 г., а количество предотвращенных ОБС в I квартале 2024 г. выросло на 59,4%, при этом объем предотвращенных ОБС за это же время вырос на 41,4%. Основным типом компьютерных атак выступает социальная инженерия (табл. 2).

Таблица 2

Динамика основных типов компьютерных атак

Тип атаки	Количество атак		Темп прироста, %
	в 2023 г. (в среднем за квартал)	в I квартале 2024 г.	
Использование методов социальной инженерии	25 669	29 477	+14,84
Фишинговые атаки	1 363	959	–29,64
Атаки с использованием вредоносного программного обеспечения	88	42	–52,27
Атаки типа «отказ в обслуживании» (DDoS)	105	108	+2,86
Иные атаки	64	70	+9,38

Примечание. Табл. 2 составлена автором по данным [4].

Нельзя однозначно сопоставить социальную инженерию как тип атаки на средства физических лиц с ее квалификацией в качестве мошеннических действий, так как для такого сопоставления на уровне самостоятельных действий клиента до непосредственного факта хищения средств нет соответствующей правовой рамки. По результатам завершенных действий с использованием социальной инженерии нет официальной статистики, в том числе у правоохранительных органов, которые не выделяют эту категорию преступлений

как самостоятельную. В условиях отсутствия у банков реальных полномочий по противодействию совершения клиентом транзакций под влиянием психологического воздействия мошенников логично, что основным инструментом борьбы с социальной инженерией в течение длительного времени выступало выделение подозрительных операций на основе анализа банковских транзакций с применением моделей машинного обучения.

Алгоритмы машинного обучения эффективнее традиционных методов, поскольку являются более адаптивными и быстрее обрабатывают массивные объемы данных, что позволяет им обнаруживать несанкционированные транзакции с достаточно высокой точностью при неизменном способе воздействия социальной инженерии. Для проверки поставленной гипотезы были разработаны модели, основанные на алгоритмах анализа больших данных. Были сформированы наиболее зарекомендовавшие себя модели машинного обучения, произведена классификация и кластеризация данных и учтены различные факторы для автоматического анализа платежных данных и выявления подозрительных транзакций.

В программной среде разработки и выполнения программного кода на языке Python в облаке – Google Colab – были импортированы библиотеки pycaret, missingno, pandas, numpy, matplotlib, seaborn, sklearn, umap-learn, holoviews и модуль scipy.stats. Для проверки гипотезы использовался набор данных Bank Account Fraud Dataset Suite (NeurIPS 2022). По заявлению разработчиков, этот набор данных:

– реалистичный, основанный на современном наборе реальных данных для обнаружения мошенничества;

– смещенный, каждый набор данных имеет отдельные контролируемые типы смещения;

– несбалансированный, этот параметр демонстрирует чрезвычайно низкую распространенность положительного класса;

– динамический, с временными данными и наблюдаемыми сдвигами распределения.

В ходе препроцессинга установлено, что набор данных NeurIPS 2022 (далее – набор данных) дисбалансирован и содержит данные о 98,9% немошеннических и 1,1% мошеннических транзакциях. В наборе данных – 1 000 000 записей о банковских транзакциях, содержащих сведения о 32 характеристиках каждой записи, пропусков данных не имеется. Записи содержат данные следующих типов: float (9), int (18), object (5). Из набора данных выделены признаки, потенциально способные стать маркером для мошенников, использующих средства социальной инженерии: доход клиента, его возраст, кредитный рейтинг и наличие других карт. Для оценки сдвигов распределения во времени используется месяц совершения транзакции. В табл. 3 приведены описательные характеристики данных в усеченном наборе.

В усеченном наборе данных имеются категориальные признаки. В целях сравнения методов машинного обучения они были преобразованы в числовые значения в ходе процесса one-hot encoding.

В отдельных признаках данных была выявлена асимметрия. Наиболее важный для текущего исследования признак fraud_bool потребовал логарифмического преобразования для устранения асимметрии в его распределении

перед применением методов, предполагающих нормальность распределения. Данные наблюдения о наличии смещений по этому признаку и отсутствию значимого смещения по другим подтверждаются графиками Q-Q Plots (рис. 1) и рассчитанными характеристиками асимметрии для признаков, отобранных для усеченного набора данных.

Таблица 3

Усеченный набор данных о банковских транзакциях

Показатели	Наименование признака данных					
	fraud_bool	**income**	**customer_age**	**credit_risk_score**	**has_other_cards**	**month**
Количество записей	1 000 000	1 000 000	1 000 000	1 000 000	1 000 000	1 000 000
Среднее значение	0,0110	0,5627	33,6891	130,9896	0,22308	3,2887
Стандартное отклонение	0,1044	0,2903	12,0258	69,6818	0,4163	2,2100
Минимальное значение	0,0000	0,1000	10,0000	–170,0000	0,0000	0,0000
25%	0,0000	0,3000	20,0000	83,0000	0,0000	1,0000
50%	0,0000	0,6000	30,0000	122,0000	0,0000	3,0000
75%	0,0000	0,8000	40,0000	178,0000	0,0000	5,0000
Максимальное значение	1,0000	0,9000	90,0000	389,0000	1,0000	7,0000
Коэффициент асимметрии	9,5208	–0,3913	0,4734	0,2795	1,3193	0,1125

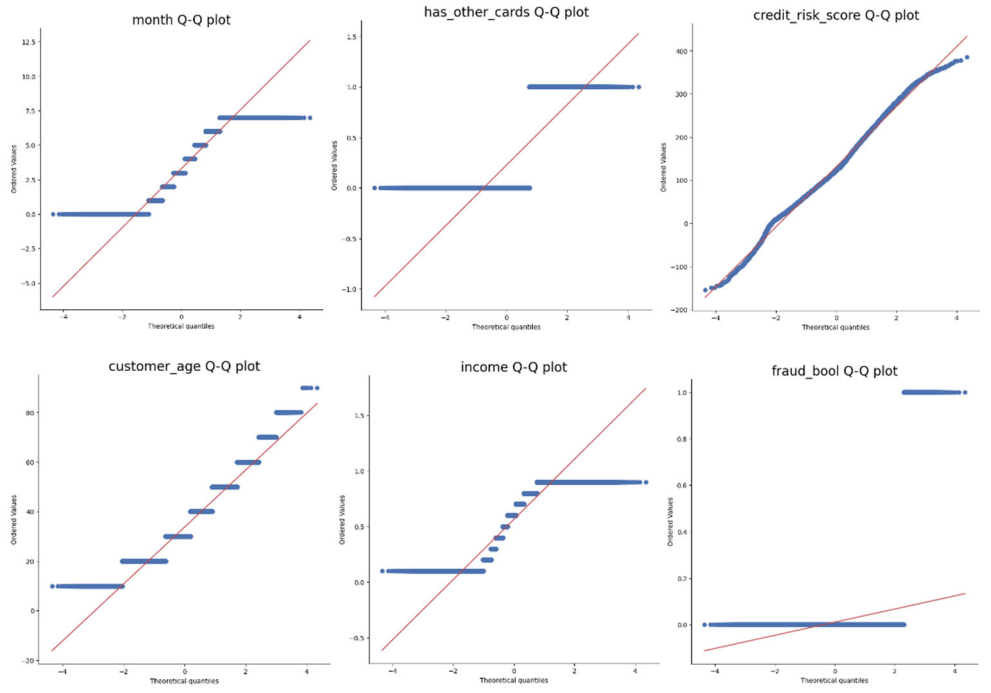

Рис. 1. Графики Q-Q для отдельных признаков данных

Корреляционный тест Пирсона на мультиколлинеарность не выявил корреляции признака мошеннических действий с ранее выделенными признаками подверженности социальной инженерии (рис. 2).

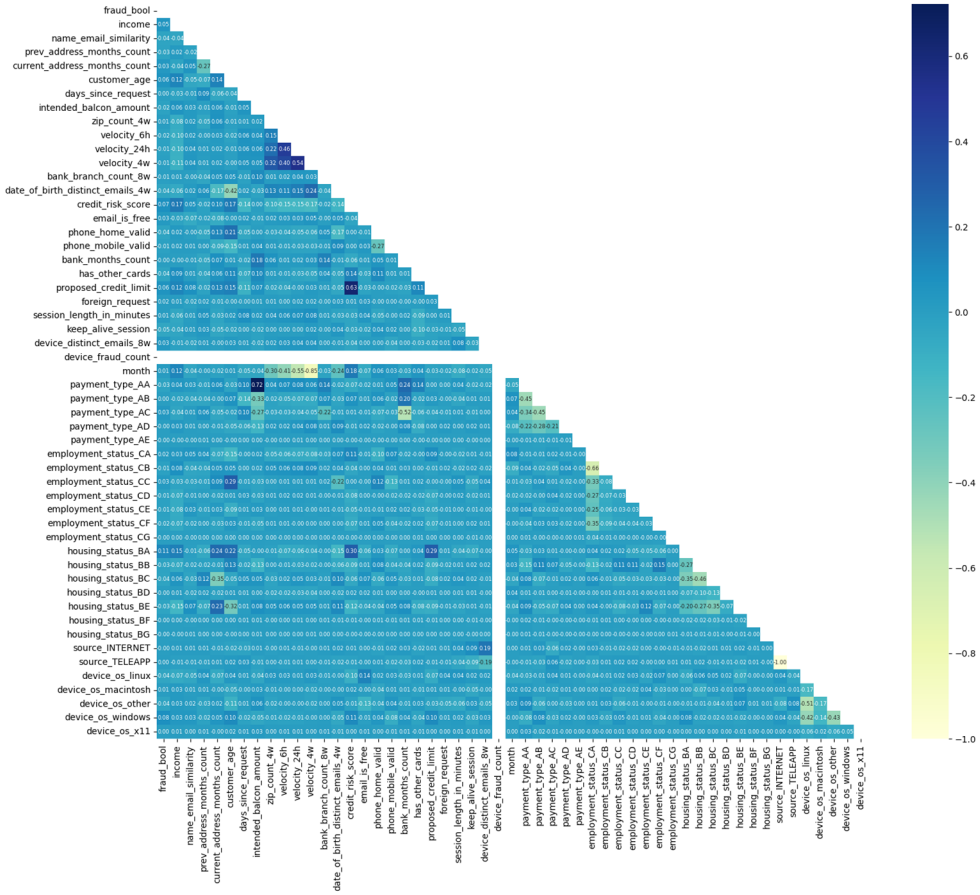

Рис. 2. Корреляционный тест Пирсона на мультиколлинеарность

Диаграмма рассеяния также не выявила однозначных закономерностей концентрации мошеннических транзакций в сочетании признаков возраста клиента и его кредитного рейтинга по исходному набору данных (рис. 3).

Выполнение алгоритма SMOTE выровняло смещение в наборе данных по признаку fraud_bool и разделило мошеннические и немошеннические операции в соотношении 50/50. На рис. 4 построена диаграмма рассеяния и раскрашена по целевому признаку – мошенническим операциям – в преобразованных данных.

Для формирования низкоразмерного вложения, которое сохранило бы основную топологическую структуру многообразия набора данных, использован алгоритм нелинейного уменьшения размерности UMAP. Для получения воспроизводимого результата начальное число random_state указано как 0.

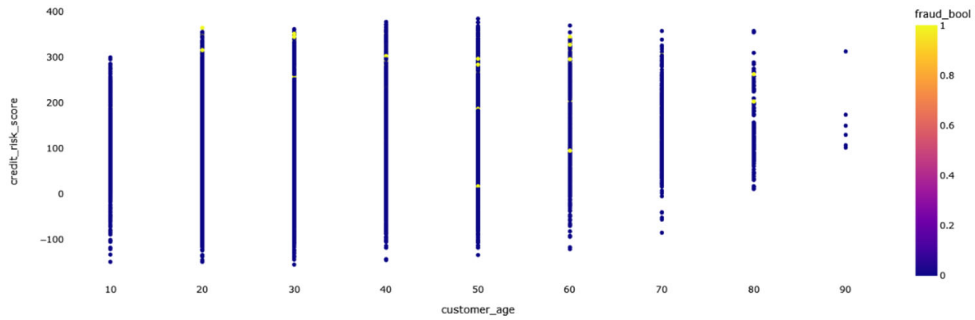

Рис. 3. Диаграмма рассеяния мошеннических операций
в сочетании признаков возраста клиента и его кредитного рейтинга

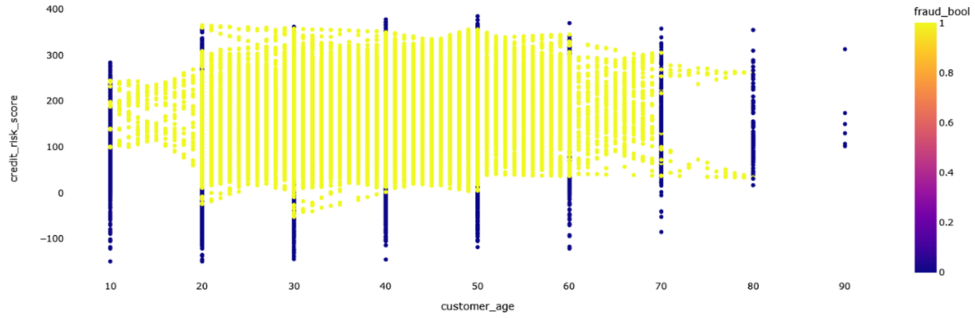

Рис. 4. Диаграмма рассеяния мошеннических операций после UMAP-обработки

Однако построенное облако точек не разделяет набор данных на классы (рис. 5).

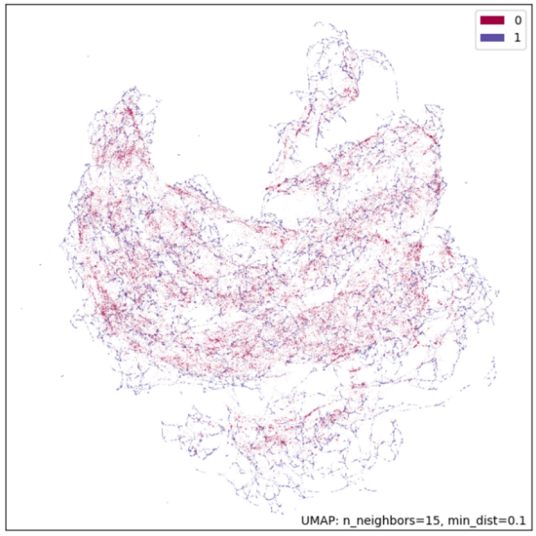

Рис. 5. Облако точек разделения мошеннических (1)
и немошеннических (0) операций

Таким образом, после применения метода SMOTE проблема классового дисбаланса, присутствующая в наборе данных, была эффективно решена, однако метод не позволяет достаточно достоверно разделить мошеннические и немошеннические операции в зависимости от каких-либо характеристик клиента или самой транзакции. В связи с этим были апробированы три модели машинного обучения, наиболее часто использующиеся в банкинге для обработки данных о транзакциях. В табл. 4 приведены характеристики моделей машинного обучения, обученных на 70% записей набора данных, преобразованных с помощью метода SMOTE.

Таблица 4

Апробация моделей машинного обучения

Модель	Метрика						
	Accuracy	**AUC**	**Recall**	**Precision**	**F1**	**Kappa**	**MCC**
Деревья решений	0,9831	0,9831	0,9869	0,9795	0,9832	0,9663	0,9663
Градиентный бустинг	0,9798	0,9973	0,9773	0,9822	0,9797	0,9595	0,9596
Логистическая регрессия	0,8063	0,8838	0,7980	0,8116	0,8047	0,6126	0,6127

Точность модели на основе деревьев решений достаточно высока и в абсолютном отношении, и по сравнению с двумя другими моделями, однако эффективность и результативность дальнейшего исследования неизбежно снижается из-за дрейфа концепций, вызывающего увеличение вероятности ошибочной классификации [16]. Структура данных о транзакциях и их взаимосвязи меняются со временем по следующим основным причинам:

– ковариантный сдвиг (сдвиг независимых переменных);

– сдвиг априорной вероятности (сдвиг целевой переменной);

– концептуальный дрейф (сдвиг во взаимоотношениях между независимой и целевой переменной).

Общий способ отслеживания дрейфа концепций заключается в том, что если показатель производительности снижается ниже порогового значения, срабатывает сигнал тревоги для повторного обучения модели [17].

В случае отслеживания социальной инженерии проблема дрейфа концепций исключительно актуальна. В 2022 г. атакам подвергались граждане старшего пенсионного возраста, которым сообщали о родственниках, попавших в автомобильные аварии. В 2024 г. все чаще фиксируются случаи мошеннических схем, в которых применяются в совокупности техники социальной инженерии и deepfake [6].

В случае дрейфа концепции модель становится переобученной и плохо работает на тестовых данных. Для регуляризации может использоваться ранняя остановка, когда машинное обучение должно остановиться в точке, где потери и ошибки в проверочных данных достигают минимального значения. Чтобы избежать ситуации недообучения, необходимо использовать планировщики скорости обучения и отслеживать параметры обратных вызовов, что реализовано в методах библиотек PyTorch, Scikit-learn и TensorFlow [21]. Поскольку в банковских антифрод-системах используется реестр (ансамбль) моделей, а не отдельные разновидности, то процесс ранней остановки моделей для переобучения можно представить в следующем виде (рис. 6).

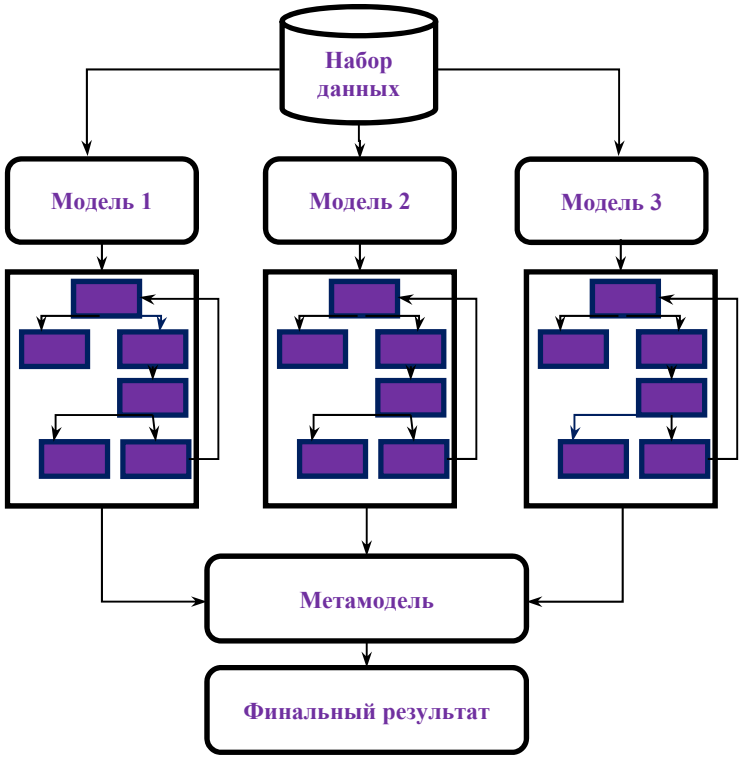

Рис. 6. Ансамблирование в метамодель алгоритмов с реализованной ранней остановкой

Итоговый ансамбль объединяет прогнозы нескольких моделей. К примеру, случайные леса, в которых используется ансамбль деревьев решений, могут обеспечить надежные прогнозы за счет уменьшения дисперсии, наблюдаемой в отдельных деревьях. Результаты базовых алгоритмов объединяются в один с помощью обучаемой метамодели, которая кроме мета-факторов может учитывать и признаки из исходного набора данных.

Выводы. В результате исследования установлено, что современные модели машинного обучения за счет алгоритмов предварительной обработки наборов данных способны достигать высокой точности обучения и обработки тестовой выборки о фактах применения социальной инженерии. Для повышения точности предсказания моделей могут использоваться ансамбли алгоритмов, в которых реализована ранняя остановка обучения. Однако в дополнение к банковским антифрод-системам требуется вмешательство регулятора, который отслеживает видоизменение социальной инженерии и использование новых методов воздействия на клиентов коммерческих банков. Ввиду значительного влияния дрейфа концепций на реальную эффективность моделей машинного обучения на практике Банк России в своей регуляторной стратегии делает ставку на формирование базы данных «О случаях и попытках осуществления переводов денежных средств без добровольного согласия клиента» на основе сведений, которые предоставляют банки и другие участники информационного обмена,

в том числе МВД России. Поскольку банки обязаны противодействовать переводам, которые происходят без добровольного согласия клиента, в частности под воздействием злоумышленников, антифрод-системы банков-отправителей, а с 1 августа 2024 г. – и банков-получателей платежей, должны регулярно переобучаться. Кроме факта фиксации в базах банковские антифрод-системы должны реагировать на нетипичный характер операции и иметь возможность отслеживать подозрительную активность клиента незадолго до операции (многочисленные телефонные звонки и СМС-сообщения, попытки входа в мобильное приложение с новых или подозрительных устройств). Точность обучения моделей по-прежнему важна, так как по требованиям законодательства банки обязаны задерживать мошеннические переводы на два дня, в противном случае им придется вернуть клиенту похищенные деньги.

Литература

1. *Аркадьева О.Г., Березина Н.В.* Формирование модели государственного регулирования развития технологий искусственного интеллекта в финансовом секторе // Oeconomia et Jus. 2023. № 4. С. 12-21. DOI: 10.47026/2499-9636-2023-4-12-21.

2. *Бердышев А.В., Зархин И.Е., Катышева А.А.* Оценка технологических возможностей противодействия мошенническим практикам в банковском секторе // Вестник университета. 2022. № 10. С. 193–204.

3. *Йоллыев А.Б.* Безопасность в банковской сфере: ключевые аспекты и роль кибербезопасности в эпоху цифровой экономики // Российский журнал менеджмента. 2024. Т. 2, № 1(70). С. 140–142.

4. Обзор отчетности об инцидентах информационной безопасности при переводе денежных средств [Электронный ресурс] // Банк России: офиц. сайт. URL: https://cbr.ru/statistics/ib/review_1q_2024/ (дата обращения: 04.09.2024).

5. *Петрякова Л.А.* Предупреждение мошенничеств в банковской сфере // Всероссийский криминологический журнал. 2023. Т. 17, № 4. С. 383–391.

6. Социальная инженерия [Электронный ресурс] // Tadviser. Государство. Бизнес. Технологии: сайт. URL: https://www.tadviser.ru/index.php (дата обращения: 04.09.2024).

7. *Федосенко М.Ю.* Разработка модели поведения злоумышленника, осуществляющего действия по легализации доходов, применительно к автоматизированным банковским системам дистанционного обслуживания // Экономика и качество систем связи. 2022. № 4. С. 53–61.

8. *Bin S.R., Schetinin V., Sant P.* Review of Machine Learning Approach on Credit Card Fraud Detection. *Hum-Cent Intell Syst*, 2022, no. 2, pp. 55–68. DOI: 10.1007/s44230-022-00004-0.

9. *Boulieris P., Pavlopoulos J., Xenos A. et al.* Fraud detection with natural language processing. *Mach Learn*, vol. 113, pp. 5087–5108. DOI: 10.1007/s10994-023-06354-5.

10. *Dang T.K., Ha T.A.* Comprehensive Fraud Detection for Credit Card. Transactions in Federated Averaging. *Sn Comput Sci*, 2024, no. 5, p. 578. DOI: 10.1007/s42979-024-02898-y.

11. *G R J., P A.I.* Attention layer integrated BiLSTM for financial fraud prediction. *Multimed Tools Appl*, 2024, pp. 1–17. DOI: 10.1007/s11042-024-18764-1.

12. *Gorle V.L.N., Panigrahi S.* A semi-supervised Anti-Fraud model based on integrated XGBoost and BiGRU with self-attention network: an application to internet loan fraud detection. *Multimed Tools Appl*, 2024, no 83, pp. 56939–56964. DOI: 10.1007/s11042-023-17681-z.

13. *Hajek P., Abedin M.Z., Sivarajah U.* Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework. *Inf Syst Front*, 2023, no. 25, pp. 1985–2003. DOI: 10.1007/s10796-022-10346-6.

14. *Karbasiyan M., Hamidi H., Srinivasa R.K.* Presenting a Model to Detect the Fraud in Banking using Smart Enabling Tools. *International Journal of Engineering*, 2024, no. 37(03), pp. 529-537. DOI: 10.5829/ije.2024.37.03c.10.

15. *Lin Y.-F., Wang C.-W., Wu C.-W.* Application of Machine Learning in Credit Card Fraud Detection: A Case Study of F Bank. HCI in Business, Government and Organizations. In: HCII 2024.

Lecture Notes in Computer Science. Springer, Cham, 2024, vol. 14720, pp. 210–222. DOI: 10.1007/-978-3-031-61315-9_15.

16. *Mwiti D.* Random Forest Regression: When Does It Fail and Why? Available at: https://neptune.ai/blog/random-forest-regression-when-does-it-fail-and-why (accessed 04.08.2024).

17. *Nidhi M.V., Gupta V., Vig R.* Methods to Investigate Concept Drift in Big Data Streams. In: Knowledge Computing and Its Applications. Springer, Singapore, 2018, pp. 51–74. DOI: 10.1007/978-981-10-6680-1_3.

18. *Rutskiy V. et al.* Prospects for the Use of Artificial Intelligence to Combat Fraud in Bank Payments. Data Science and Algorithms in Systems. In: CoMeSySo 2022. Lecture Notes in Networks and Systems. Springer, Cham, 2023, vol. 597, pp. 959-971.

19. *Salam A.M., Fouad K.M., Elbably D.L. et al.* Federated learning model for credit card fraud detection with data balancing techniques. *Neural Comput & Applic*., 2024, no. 36, pp. 6231–6256. DOI: 10.1007/s00521-023-09410-2.

20. *Sergadeeva A.I., Lavrova D.S., Zegzhda D.P.* Bank Fraud Detection with Graph Neural Networks. *Aut. Control Comp. Sci*, 2022, no. 56, pp. 865–873. DOI: 10.3103/S0146411622080223.

21. *Shinde S.D.* Pause for Performance: The Guide to Using Early Stopping in ML and DL Model Training. Available at: https://pub.towardsai.net/pause-for-performance-the-guide-to-using-early-stopping-in-ml-and-dl-model-training-0abd24e5cdcb (accessed 04.09.2024).

22. *Vanini P., Rossi S., Zvizdic E. et al.* Online payment fraud: from anomaly detection to risk management. *Financ Innov*, 2023, no. 9, pp. 66. DOI: 10.1186/s40854-023-00470-w.

23. *Vashistha A., Tiwari A.K.* Building Resilience in Banking Against Fraud with Hyper Ensemble Machine Learning and Anomaly Detection Strategies. *Sn Comput sci*, 2024, no. 5, p. 556. DOI: 10.1007/s42979-024-02854-w.

24. *Vashistha A., Tiwari A.K., Singh P. et al.* A Robust Framework for fraud Detection in Banking using ML and NN. Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci. 2024, no. 94, pp. 201–212. DOI: 10.1007/s40010-024-00871-1.

25. *Zioviris G., Kolomvatsos K., Stamoulis G.* An intelligent sequential fraud detection model based on deep learning. *J Supercomput*, 2024, no. 80, p. 14824–14847. DOI: 10.1007/s11227-024-06030-y.

АРКАДЬЕВА ОЛЬГА ГЕННАДЬЕВНА – кандидат экономических наук, доцент кафедры финансов, кредита и экономической безопасности, Чувашский государственный университет, Россия, Чебоксары (knedlix@yandex.ru; ORCID: https://orcid.org/0000-0003-4868-2365).

Olga G. ARKADEVA

THE USE OF MACHINE LEARNING TECHNIQUES TO TRACK SOCIAL ENGINEERING IN BANKING TRANSACTIONS

Key words: *data imbalance, asymmetry, oversampling, concept drift, early stop, machine learning models.*

The relevance of the study is due to the fact that in the era of digitalization and a widespread introduction of technologies, online payments and other remote banking transactions are becoming increasingly popular, which leads to an increase in cases of fraud using social engineering.

The purpose of the study *is to assess the opportunities of machine learning methods to counteract social engineering, as well as to identify limitations of using these methods in banking anti–fraud systems.*

Materials and methods*. To test the hypothesis of insufficient effectiveness of machine learning methods for recognizing attacks using social engineering, the Bank Account Fraud Dataset Suite (NeurIPS 2022) was used. Data preparation for machine learning was implemented using UMAP nonlinear dimensionality reduction algorithm and SMOTE synthetic oversampling method.*

Study results. *Clustering of data did not make it possible to reliably separate fraudulent and non-fraudulent transactions depending on any characteristics of the client or the transaction itself. At this, the accuracy of the model based on decision tree is the highest compared to gradient boosting and logistic regression, however, the efficiency and effectiveness of practical using the algorithms inevitably decreases due to the drift of concepts, which causes an increase in the probability of erroneous classification. In case of concept drift, the model becomes over-trained and does not work well on test data. An early stop can be used for regularization, when the model learning process should stop at the point where losses and errors in the verification data reach a minimum value.*

Conclusions. *Since banks are obliged to counteract transfers that occur without the client's voluntary consent, in particular under the influence of intruders, the anti-fraud systems of sending and receiving banks should be regularly retrained in order to avoid the spread of fraudulent cases and banks receiving losses from the need to make compensation payments. The practical significance lies in the possibility of using the results to track and counteract social engineering in banking transactions in order to further improve regulatory efforts and actions of commercial banks themselves.*

References

1. Arkad'eva O.G., Berezina N.V. *Formirovanie modeli gosudarstvennogo regulirovaniya razvitiya tekhnologii iskusstvennogo intellekta v finansovom sektore* [Formation of a model of state regulation of the development of artificial intelligence technologies in the financial sector]. *Oeconomia et Jus,* 2023, no. 4, pp. 12–21. DOI: 10.47026/2499-9636-2023-4-12-21.

2. Berdyshev A.V., Zarkhin I.E., Katysheva A.A. *Otsenka tekhnologicheskikh vozmozhnostei protivodeistviya moshennicheskim praktikam v bankovskom sektore* [Assessing Technological Capabilities to Combat Fraudulent Practices in the Banking Sector]. *Vestnik universiteta,* 2022, no. 10, pp. 5193–204.

3. Iollyev A.B. *Bezopasnost' v bankovskoi sfere: klyuchevye aspekty i rol' kiberbez-opasnosti v epokhu tsifrovoi ekonomiki* [Security in the banking sector: key aspects and the role of cybersecurity in the era of the digital economy]. *Rossiiskii zhurnal menedzhmenta*, 2024, vol. 2, no. 1(70), pp. 140–142.

4. *Obzor otchetnosti ob intsidentakh informatsionnoi bezopasnosti pri perevode denezhnykh sredstv. Ofitsial'nyi sait Banka Rossii* [Review of reporting on information security incidents during money transfers. Official website of the Bank of Russia]. Available at: https://cbr.ru/statistics/ib/review_1q_2024/ (Access Date: 2024, Sept. 4).

5. Petryakova L.A. *Preduprezhdenie moshennichestv v bankovskoi sfere* [Prevention of fraud in the banking sector]. *Vserossiiskii kriminologicheskii zhurnal*, 2023, vol. 17, no. 4, pp. 383–391.

6. *Sotsial'naya inzheneriya. Tadviser. Gosudarstvo. Biznes. Tekhnologii*. [Social Engineering. Tadviser. State. Business. Technologies]. Available at: https://www.tadviser.ru/index.php (Access Date: 2024, Sept. 4).

7. Fedosenko M.Yu. *Razrabotka modeli povedeniya zloumyshlennika, osushchestvlyayushchego deistviya po legalizatsii dokhodov, primenitel'no k avtomatizirovannym bankovskim sistemam distantsionnogo obsluzhivaniya* [Development of a behavior model for an attacker carrying out actions to legalize income, as applied to automated remote banking systems]. *Mezhdunarodnyi nauchnyi zhurnal «Ekonomika i kachestvo sistem svyazi»*, 2022, no 4, pp. 53-61.

8. Bin S.R., Schetinin V., Sant P. Review of Machine Learning Approach on Credit Card Fraud Detection. *Hum-Cent Intell Syst*, 2022, no. 2, pp. 55–68. DOI: 10.1007/s44230-022-00004-0.

9. Boulieris P., Pavlopoulos J., Xenos A. et al. Fraud detection with natural language processing. *Mach Learn*, vol. 113, pp. 5087–5108. DOI: 10.1007/s10994-023-06354-5.

10. Dang T.K., Ha T.A. Comprehensive Fraud Detection for Credit Card. Transactions in Federated Averaging. *Sn Comput Sci*, 2024, no. 5, p. 578. DOI: 10.1007/s42979-024-02898-y.

11. G R J., P A.I. Attention layer integrated BiLSTM for financial fraud prediction. *Multimed Tools Appl*, 2024, pp. 1–17. DOI: 10.1007/s11042-024-18764-1.

12. Gorle V.L.N., Panigrahi S. A semi-supervised Anti-Fraud model based on integrated XGBoost and BiGRU with self-attention network: an application to internet loan fraud detection. *Multimed Tools Appl*, 2024, no 83, pp. 56939–56964. DOI: 10.1007/s11042-023-17681-z.

13. Hajek P., Abedin M.Z., Sivarajah U. Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework. *Inf Syst Front*, 2023, no. 25, pp. 1985–2003. DOI: 10.1007/s10796-022-10346-6.

14. Karbasiyan M., Hamidi H., Srinivasa R.K. Presenting a Model to Detect the Fraud in Banking using Smart Enabling Tools. *International Journal of Engineering*, 2024, no. 37(03), pp. 529-537. DOI: 10.5829/ije.2024.37.03c.10.

15. Lin Y.-F., Wang C.-W., Wu C.-W. Application of Machine Learning in Credit Card Fraud Detection: A Case Study of F Bank. HCI in Business, Government and Organizations. In: HCII 2024. Lecture Notes in Computer Science. Springer, Cham, 2024, vol. 14720, pp. 210–222. DOI: 10.1007/978-3-031-61315-9_15.

16. Mwiti D. Random Forest Regression: When Does It Fail and Why? Available at: https://neptune.ai/blog/random-forest-regression-when-does-it-fail-and-why (accessed 04.08.2024).

17. Nidhi M.V., Gupta V., Vig R. Methods to Investigate Concept Drift in Big Data Streams. In: Knowledge Computing and Its Applications. Springer, Singapore, 2018, pp. 51–74. DOI: 10.1007/978-981-10-6680-1_3.

18. Rutskiy V. et al. Prospects for the Use of Artificial Intelligence to Combat Fraud in Bank Payments. Data Science and Algorithms in Systems. In: CoMeSySo 2022. Lecture Notes in Networks and Systems. Springer, Cham, 2023, vol. 597, pp. 959–971.

19. Salam A.M., Fouad K.M., Elbably D.L. et al. Federated learning model for credit card fraud detection with data balancing techniques. *Neural Comput & Applic*., 2024, no. 36, pp. 6231–6256. DOI: 10.1007/s00521-023-09410-2.

20. Sergadeeva A.I., Lavrova D.S., Zegzhda D.P. Bank Fraud Detection with Graph Neural Networks. *Aut. Control Comp. Sci*, 2022, no. 56, pp. 865–873. DOI: 10.3103/S0146411622080223.

21. Shinde S.D. Pause for Performance: The Guide to Using Early Stopping in ML and DL Model Training. Available at: https://pub.towardsai.net/pause-for-performance-the-guide-to-using-early-stopping-in-ml-and-dl-model-training-0abd24e5cdcb (accessed 04.09.2024).

22. Vanini P., Rossi S., Zvizdic E. et al. Online payment fraud: from anomaly detection to risk management. *Financ Innov*, 2023, no. 9, pp. 66. DOI: 10.1186/s40854-023-00470-w.

23. Vashistha A., Tiwari A.K. Building Resilience in Banking Against Fraud with Hyper Ensemble Machine Learning and Anomaly Detection Strategies. *Sn Comput sci*, 2024, no. 5, p. 556. DOI: 10.1007/s42979-024-02854-w.

24. Vashistha A., Tiwari A.K., Singh P. et al. A Robust Framework for fraud Detection in Banking using ML and NN. Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci. 2024, no. 94, pp. 201–212. DOI: 10.1007/s40010-024-00871-1.

25. Zioviris G., Kolomvatsos K., Stamoulis G. An intelligent sequential fraud detection model based on deep learning. *J Supercomput*, 2024, no. 80, p. 14824–14847. DOI: 10.1007/s11227-024-06030-y.

OLGA G. ARKADEVA – Candidate of Economics Sciences, Associate Professor, Department of Finance, Credit and Economic Security, Chuvash State University, Russia, Cheboksary (knedlix@yandex.ru; ORCID: https://orcid.org/0000-0003-4868-2365).

Формат цитирования: *Аркадьева О.Г.* Использование методов машинного обучения для отслеживания социальной инженерии в банковских транзакциях [Электронный ресурс] // Oeconomia et Jus. 2024. № 4. С. 1–14. URL: http://oecomia-et-jus.ru/single/2024/4/1. DOI: 10.47026/2499-9636-2024-4-1-14.