ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ
INFORMATION TECHNOLOGY AND TELECOMMUNICATIONS

УДК 004.89
CC BY 4.0
DOI: 10.21822/2073-6185-2024-51-4-23-32 Оригинальная статья/ Original article

Аналитическая оценка методов обнаружения мошенничества с банковскими картами: обучение с учителем, без учителя и с подкреплением

Абдурахман Джамал Джама
Финансовый университет при Правительстве РФ,
125167, Москва, пр-кт Ленинградский, д. 49/2, Россия

Резюме. Цель. Мошенничество с банковскими картами становится все более серьезной проблемой для частных лиц, предприятий и финансовых учреждений. Возникает необходимость в применении эффективных мер по обнаружению мошенничества для защиты потребителей и бизнеса от финансовых потерь. **Метод.** Применен теоретико-информационный анализ методов обнаружения мошенничества с банковскими картами и определен потенциал алгоритмов машинного обучения в повышении точности обнаружения мошенничества. **Результат.** Дана аналитическая оценка методов обнаружения мошенничества, охватывающая различные подходы к обучению, включая методы контролируемого, неконтролируемого обучения и обучения с подкреплением. **Вывод.** Выбор метода обнаружения мошенничества должен основываться на всестороннем понимании доступных данных, конкретных требований в области применения и компромисса между различными методами с точки зрения производительности, адаптируемости и вычислительной сложности.

Ключевые слова: мошенничество, банковские карты, машинное обучение, обучение с учителем, без учителя, с подкреплением, несбалансированный набор данных

Для цитирования: Абдурахман Джамал Джама. Аналитическая оценка методов обнаружения мошенничества с банковскими картами: обучение с учителем, без учителя и с подкреплением. Вестник Дагестанского государственного технического университета. Технические науки. 2024; 51(4):23-32. DOI:10.21822/2073-6185-2024-51-4-23-32

An analytical assessment of credit card fraud detection techniques: Supervised, Unsupervised, and Reinforcement Learning

Abdourahman Djamal Djama
Financial University under the Government of the Russian Federation,
49/2 Leningradsky Ave., Moscow 125167, Russia

Abstract. Objective. Bank card fraud is an increasingly serious problem for individuals, businesses and financial institutions. There is a need for effective fraud detection measures to protect consumers and businesses from financial losses. **Method.** information-theoretical analysis of methods for detecting fraud with bank cards, machine learning algorithms in improving the accuracy of fraud detection. **Result.** An analytical evaluation of fraud detection methods is provided, covering different learning approaches: supervised, unsupervised and reinforcement learning. **Conclusion.** The choice of a fraud detection method should be based on an understanding of the available data, the specific requirements of the application domain and the trade-offs between methods in terms of performance, adaptability and computational complexity.

Keywords: fraud, bank cards, machine learning, supervised learning, unsupervised learning, reinforcement learning, imbalanced dataset

For citation: Abdourahman Djamal Djama. An analytical assessment of credit card fraud detection techniques: Supervised, Unsupervised, and Reinforcement Learning. Herald of the Da-

ghestan State Technical University. Technical Sciences. 2024; 51(4):23-32. DOI:10.21822/2073-6185-2024-51-4-23-32

Введение. Мошенничество с банковскими картами по-прежнему представляет собой серьезную проблему для финансовых учреждений, вызывая значительные финансовые потери подрывая доверие потребителей. С развитием цифровых транзакций потребность в надежных системах обнаружения мошенничества стала более острой, чем когда-либо. Традиционные методы часто не справляются с обнаружением сложных моделей мошенничества, что требует изучения передовых методов машинного обучения.

Постановка задачи. В статье представлена аналитическая оценка методов обнаружения мошенничества с банковскими картами с упором на контролируемые, неконтролируемые подходы и подходы к обучению с подкреплением. Обучение с учителем использует размеченные данные для обучения моделей прогнозированию мошеннических транзакций, тогда как обучение без учителя выявляет аномалии в данных транзакций без предварительной маркировки.

С другой стороны, обучение с подкреплением динамически изучает оптимальные стратегии обнаружения посредством непрерывной обратной связи. Оценивая эти методы, мы стремимся понять их сильные стороны, ограничения и практическое применение для повышения точности и эффективности обнаружения мошенничества. Сравнительный анализ предоставит ценную информацию для разработки более эффективных систем обнаружения мошенничества, что в конечном итоге защитит финансовые транзакции в эпоху цифровых технологий.

Методы исследования. 1. Контролируемые методы обучения. Supervised learning (Обучение с учителем) - это тип парадигмы машинного обучения, в которой алгоритм обучается на помеченном наборе данных. Это означает, что каждый вход в обучающих данных связан с соответствующим выходом или целью. Цель обучения с учителем - изучить сопоставление или функцию входных данных с выходными данными, чтобы при представлении новых, невидимых входных данных алгоритм мог делать точные прогнозы или классификации [2].

1.1. Машина опорных векторов (Support Vector Machine). Машины опорных векторов (SVM) стали известным алгоритмом машинного обучения для обнаружения мошенничества с банковскими картами. В качестве метода контролируемого обучения SVM способен классифицировать данные на две категории: мошеннические и немошеннические транзакции на основе функций, извлеченных из данных транзакций. Алгоритм учится распознавать закономерности и особенности данных, которые отличают мошеннические транзакции от законных [6].

Одна из ключевых сильных сторон SVM заключается в их способности идентифицировать нелинейные границы принятия решений путем преобразования данных с помощью нелинейной функции **φ** в многомерное пространство.

Это преобразование позволяет разделять точки данных, которые не могут быть разделены прямой линией в исходном пространстве, путем создания линейной гиперплоскости в пространстве признаков **F**, эффективно разделяя точки разных классов. Впоследствии гиперплоскость проецируется обратно в исходное пространство **I**, принимая вид нелинейной кривой.

Математически, учитывая *n* выборок обучающих данных:

$$\{(x_i, y_i)\}_{i=1}^{n}, \quad x_i \in R^N, y_i \in \{-1, 1\} \tag{1}$$

SVM формулируется следующей оптимизационной задачей:

$$Minimize \quad \Phi(w) = \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i \tag{2}$$

При условии

$$y_i(\langle w, \phi(x_i)\rangle + b) \geq 1 - \xi_i, \quad i = 1, \ldots, n$$
$$\xi_i \geq 0, \quad i = 1, \ldots, n \tag{3}$$

Где, функция ядра φ отображает точки обучения x_i из входного пространства в пространство признаков более высокой размерности. Параметр регуляризации C контролирует компромисс между достижением низкой ошибки в обучающих данных и минимизацией нормы весов. Преимуществом использования SVM является его способность обрабатывать многомерные данные и нелинейные связи между функциями, что может повысить эффективность обнаружения мошенничества с банковскими картами.

1.2. Случайный лес (Random Forest) является алгоритмом машинного обучения, применяемым для задач классификации и регрессионного анализа. Этот метод широко применяется в задачах выявления мошенничества с использованием банковских карт. Его преимущество заключается в способности обрабатывать обширные наборы данных с многомерными характеристиками.

Для обнаружения мошенничества с банковскими картами модель Random Forest обучается на наборе данных транзакций, где каждая операция помечена как мошенническая или легитимная [3]. Алгоритм формирует несколько деревьев решений, каждое из которых обучается на случайной выборке данных и случайном наборе признаков.

Деревья принимают коллективное решение по предсказанию статуса новой транзакции как мошеннической или нет. Random Forest демонстрирует эффективность в задаче обнаружения мошенничества с банковскими картами за счет способности обрабатывать дисбалансированные данные, где количество мошеннических операций значительно меньше числа легитимных транзакций.

Кроме того, этот метод позволяет выявить наиболее значимые признаки, способствующие обнаружению мошенничества. Также Random Forest способен обрабатывать отсутствующие значения и выбросы, часто встречающиеся в данных по банковским картам. В контексте классификации классификатор случайного леса m получается большинством голосов среди K деревьев классификации с входом x и Θ - набор параметров, т.е:

$$m(x : \Theta_1, ..., \Theta_K) = \begin{cases} 1 & if \frac{1}{K}\sum_{j=1}^{K} m(x; \Theta_j) > \frac{1}{2} \\ 0 & otherwise \end{cases} \tag{4}$$

2. Неконтролируемые методы обучения. Обучение без учителя (Unsupervised learning) - это разновидность машинного обучения, которая занимается анализом и кластеризацией немаркированных данных. В отличие от обучения с учителем, которое связано с помеченными наборами данных для обучения моделей прогнозированию результатов, обучение без учителя работает с данными, которые не имеют помеченных ответов. Основная цель - найти основную структуру или распределение данных, что позволит обнаружить закономерности и идеи без необходимости вмешательства человека [2].

2.1. Ограниченная машина Больцмана (Restricted Boltzmann Machine).

Огранченная машина Больцмана (RBM) - это тип алгоритма глубокого обучения, который можно использовать для обнаружения мошенничества с банковскими картами. RBM является генеративная модель, которая изучает распределение вероятностей входных данных и может использоваться для обнаружения аномалий в данных.

При обнаружении мошенничества с банковскими картами RBM можно обучить на большом наборе данных о транзакциях по банковскими картам, чтобы изучить нормальные модели поведения для законных транзакций. Как только RBM изучит эти шаблоны, его можно использовать для выявления аномалий в новых транзакциях, которые не соответствуют нормальному шаблону. Эти аномалии могут свидетельствовать о мошеннической деятельности.

Модель RBM состоит из видимых и скрытых слоев, которые связаны симметричными весами. Входы *x* соответствуют нейронам в видимом слое. Реакция нейронов *h* в скрытом слое моделирует распределение вероятностей входных данных. Распределение вероятностей получается путем изучения симметричных связующих весов между видимыми и скрытыми слоями. Нейроны в одном слое не связаны.

P(h|x) - Условная вероятность конфигурации скрытых нейронов (*h*) при заданной конфигурации видимых нейронов, связанных с входами (*x*), равна:

$$p(h|x) = \prod_i p(h_i|x) \tag{5}$$

Генеративное обучение в RBM используется для итеративного изучения неизвестного (*h*) с использованием входных данных (*x*). Фаза генеративного обучения повторяется до тех пор, пока реконструированные образцы не будут максимально приближены к *x.*

Кроме того, RBM используется для предварительной обработки данных и извлечения соответствующих функций, которые затем можно использовать в другом алгоритме машинного обучения, таком как дерево решений или нейронная сеть.

2.2. Автоэнкодер (Auto-Encoder) представляет собой вид нейронной сети, который применяется для обнаружения мошенничества с использованием банковских карт. Суть использования автоэнкодеров заключается в выявлении аномалий в данных [10]. Эти модели обучаются восстанавливать входные данные, и любое отклонение от ожидаемого результата рассматривается как аномалия. AE определяется по следующей формуле:

$$\hat{X} = D(E(X)) \tag{6}$$

Где, X - входные данные, D - карта декодирования, E - карта кодирования, а $\hat{X}$ - реконструированные входные данные. Задача автоэнкодера - максимально точно аппроксимировать распределение X.

Для обнаружения мошенничества с банковскими картами часто применяются автоэнкодеры. В качестве входных данных для автоэнкодера выступают информация о транзакциях, такая как сумма транзакции, местоположение, время и другие. Автоэнкодер обучается восстанавливать эту информацию, и любое отклонение от ожидаемого вывода идентифицируется как потенциальное мошенничество [17].

Кроме того, автоэнкодеры могут быть использованы и в режиме обучения без учителя. В таких случаях модель обучается на наборе данных, в котором отсутствуют мошеннические транзакции. После завершения обучения автоэнкодера, он может применяться для обнаружения любых новых транзакций, которые не соответствуют изученному распределению.

3. Обучение с подкреплением (Reinforcement learning) - это тип машинного обучения, при котором агент учится принимать решения, взаимодействуя с окружающей средой и получая обратную связь (feedback) в виде вознаграждений или наказаний в зависимости от своих действий. Этот итерационный процесс включает в себя метод проб и ошибок (trial-and-error approach), при котором стратегия агента со временем уточняется за счет максимизации совокупного вознаграждения [16]. Ключевые компоненты обучения с подкреплением включают агента, среду, действия, состояния и вознаграждения.

3.1.Алгоритм глубокой Q-сети (Deep Q Network (DQN)). Одним из таких алгоритмов обучения с подкреплением является Deep Q-Network (DQN), который сочетает в себе глубокие нейронные сети (DNN) и Q-learning для решения сложных задач принятия решений.

Q-learning - это метод обучения вне политики, целью которого является изучение оптимальной политики для агента путем оценки Q-value. Q-value пары состояние-действие (s, a), обозначаемое как Q(s, a), представляет собой ожидаемую совокупную будущую награду за выполнение действия «a» в состоянии «s» и последующее следование

оптимальной политике. Глубокая нейронная сеть используется для аппроксимации функции Q-значения, обозначаемой как Q(s, a; θ), где θ представляет параметры сети.

Входными данными для сети является состояние «s», и она выводит значения Q для всех возможных действий. Сеть обучена минимизировать разницу между прогнозируемыми значениями Q и целевыми значениями Q, которые рассчитываются с использованием уравнения Беллмана и вознаграждений, полученных от окружающей среды [8].

Deep Q Network (DQN) можно применить для обнаружения мошенничества с банковскими картами, определив проблему обнаружения мошенничества как задачу обучения с подкреплением. Пространство состояний *(state)* ***S*** включает в себя характеристики транзакций по банковским картам, обозначаемые как ***s***, такие как сумма, продавец, время, и т.д.

Пространство действий *(action)* ***A*** состоит из возможных действий, которые агент может предпринять для каждой транзакции, например одобрение, отклонение или отметка. Q-функция, ***Q(s,a)***, оценивает ожидаемое совокупное вознаграждение за выполнение действия *a* в состоянии ***s***.

Цель DQN - изучить оптимальную Q-функцию, которая максимизирует совокупное вознаграждение с течением времени.

$$Q(s,a) = E\left[\sum_{t=0}^{\infty} \gamma^t r_t \mid s_0 = s, a_0 = a\right] \quad (7)$$

Где, *rt* - награда на временном шаге «t»; γ - это коэффициент дисконтирования, который придает большее значение немедленному вознаграждению, чем будущему вознаграждению. В процессе обучения, DQN минимизирует TDE (Temporal Difference Error), δ, определяемую формулой:

$$\delta = (r + \gamma \max_{a'} Q(s', a') - Q(s, a)) \quad (8)$$

Здесь, ***r*** - непосредственный вознаграждение, ***γ*** - коэффициент дисконтирования, ***s'*** - следующее состояние, **a'** - следующее действие. DQN учится на прошлом опыте, используя воспроизведение опыта, сохранение и случайную выборку переходов (*s,a,r,s'*).

Веса Q-сети обновляются, чтобы минимизировать MSTDE (mean squared temporal difference error). Deep Q Network оценивает Q-value для текущего состояния транзакции ***s*** и выбирает действие с самым высоким Q-value или следует стохастической политике, основанной на Q-values.

Непрерывное обучение облегчается за счет периодического обновления DQN новыми данными о транзакциях, что обеспечивает адаптируемость к развивающимся моделям мошенничества. В целом, платформа DQN предлагает Data-driven и динамический подход к обнаружению мошенничества с банковскими картами, используя принципы обучения с подкреплением [19].

Обсуждение результатов. Набор данных и предварительная обработка. Этот набор данных содержит транзакции по банковскими картам, совершенные в сентябре 2019 года держателями карт. Транзакции, включают 570 записи о мошенничестве из 300 807 транзакций.

Набор данных сильно несбалансирован, и мошеннический класс составляет 0,17% всех транзакций. Из-за проблемы конфиденциальности, набор данных содержит входные переменные, которые получены из преобразования PCA. Для нечисловых характеристик «Время» и «Сумма» мы нормализуем их с помощью RobustScaler, который масштабирует данные в соответствии с диапазоном квантилей.

Специально для моделей обучения с учителем, чтобы решить проблему сильной несбалансированности, используется Random Downsampling, чтобы избежать смещения результатов в сторону немошеннического класса. В результате Random Downsampling, немошеннические транзакции случайным образом сокращаются до той же суммы, что и мошеннические транзакции.

Метрики оценки. В зависимости от цели эксперимента, который мы проводим, мы можем использовать различные статистические показатели для оценки бинарных классификаций.

Accuracy (точность) является широко используемым показателем в машинном обучении, но она может быть не самым подходящим показателем для обнаружения мошенничества с банковскими картами, особенно при работе с несбалансированными наборами данных, где количество немошеннических транзакций значительно превышает количество мошеннических транзакций. В таких случаях точность может вводить в заблуждение, поскольку модель может достичь высокой точности, просто предсказывая класс большинства (без мошенничества) без эффективного выявления мошеннических транзакций.

В данной статье рассматриваем Accuracy наряду с другим показателем, таким как площадь под кривой ROC (AUC-ROC), чтобы обеспечить более полную оценку эффективности модели, особенно в контексте обнаружения мошенничества с банковскими картами.

Accuracy (точность) - показывает, сколько правильных предсказаний сделала модель относительно общего числа предсказаний.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{9}$$

F1-Score - это показатель, который уравновешивает Precision и Recall, что делает его особенно полезным в ситуациях, когда существует дисбаланс между классами, как это часто бывает при обнаружении мошенничества с банковскими картами. F1-score - это гармоническое среднее значение точности и полноты, предоставляющее единое значение, учитывающее как ложноположительные, так и ложноотрицательные результаты. Формула расчета F1-Score выглядит следующим образом:

$$\text{F1 Score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{10}$$

Площадь под кривой операционной характеристики приемника (AUC-ROC) объединяет уровень ложных срабатываний (FPR) и уровень истинного положительного результата (*TPR*) в единый показатель, рассматривая класс мошенничества как «положительный», а класс отсутствия мошенничества как «отрицательный». $TPR = TP/P$ и $FPR = FP/N$, где P и N обозначают количество образцов положительного и отрицательного классов соответственно.

TP (истинно положительный) - это количество корректно предсказанных положительных образцов, а *FP* (ложно положительный) - количество неверно предсказанных положительных образцов. AUC-ROC оценивает производительность модели при различных пороговых значениях принятия решений, поэтому он очень полезен в случае несбалансированных наборов данных, так как он.

Это помогает найти оптимальный баланс между ложноположительными и ложноотрицательными результатами, учитывая особенности системы обнаружения мошенничества с банковскими картами.

Для оценки эффективности обнаружения мошеннических транзакций был использован метод *k*-кратной кросс-валидации для предотвращения переобучения. Процесс включает разделение набора данных на *k* равных частей или сгибов.

Одна часть используется как тестовый набор данных, а остальные (*k-1*) части - для обучения модели. Этот процесс повторяется *k* раз, каждый раз используя один из частей в качестве тестового набора данных. Затем результаты *k* итераций усредняются для получения оценки производительности модели. Кросс-валидация *k-Fold* обеспечивает

более точную оценку производительности модели на независимых данных, тестируя модель на нескольких подмножествах данных.

В табл. 1 представлены результаты проведенного сравнительного анализа методов обнаружения мошенничества с банковскими картами в различных парадигмах машинного обучения: обучение с учителем (SVM, RF), обучение без учителя (RBM, AE) и обучение с подкреплением (DQN).

Таблица 1. Результаты проведенного сравнительного анализа методов обнаружения мошенничества

Table 1. Results of the comparative analysis of fraud detection methods

Learning Algorithms	Accuracy	F1-Score	AUC-ROC
Support vector machine (SVM)	0.983	0.911	0.979
Random Forest (RF)	0.992	0.957	0.987
Restricted Boltzmann Machine (RBM)	0,978	0.881	0.952
Auto-Encoder (AE)	0.987	0.934	0.936
Deep Q Network (DQN)	0.978	0.926	0.965

Метрики оценки включают Accuracy, показатель F1-score и площадь под кривой рабочей характеристики приемника (AUC-ROC).

Для контролируемого обучения мы использовали модели случайного леса и машины опорных векторов (SVM). Модель Random Forest превзошла другие с точностью 99,26%, показателем F1-score 95,70% и AUC-ROC 98,75%. Модель SVM достигла точности 98,3%, показателя F1-score 91,06% и AUC-ROC 97,9%. Эти результаты показывают, что модели контролируемого обучения, особенно случайный лес, очень эффективны в обнаружении мошенничества с банковскими картами.

Модели обучения без учителя, включая кластеризацию. Ограниченная машина Больцмана (Restricted Boltzmann Machine (RBM)) и Автоэнкодер (Auto-Encoder), показали более низкую производительность по сравнению с моделями с учителем. Модель кластеризации RBM имела точность 97,82%, показатель F1-score 88,05% и AUC-ROC 95,25%. Модель Автоэнкодер показала лучшие результаты: точность 98,70%, показатель F1-score 93,42% и AUC-ROC 88,30%.

Хотя эти модели полезны в сценариях, в которых отсутствуют размеченные данные, их более высокий уровень ложных срабатываний подчеркивает необходимость дальнейшего совершенствования, чтобы стать жизнеспособной альтернативой контролируемому обучению. Обучение с подкреплением было представлено моделью Deep Q-Network (DQN), которая дала многообещающие результаты. Модель DQN зафиксировала точность 97,85%, показатель F1 92,66% и AUC-ROC 96,55%. Хотя его производительность была немного ниже, чем у наиболее эффективных контролируемых моделей, она значительно превосходила неконтролируемые модели.

Подход с подкреплением обучения демонстрирует потенциал адаптации в динамичной среде, где модели мошенничества постоянно развиваются. Изучение различных методов обнаружения мошенничества с банковскими картами в рамках контролируемых, неконтролируемых моделей обучения и моделей обучения с подкреплением обеспечивает полное понимание их сильных и слабых сторон. Модели контролируемого обучения продемонстрировали исключительную производительность, особенно при работе с хорошо размеченными данными, что подчеркивает их эффективность в таких сценариях.

Высокая точность, показатель F1 и значения AUC-ROC, достигнутые с помощью этих моделей, указывают на их способность обнаруживать мошенничество при наличии достаточного количества размеченных данных. Среди методов контролируемого обучения выдающейся моделью стал случайный лес, способный выявлять сложные закономерности и взаимодействия в данных транзакций, а машины опорных векторов (SVM) также продемонстрировали похвальную производительность.

Напротив, модели обучения без учителя, такие как ограниченные машины Больцмана и кластеризация с автоматическим кодированием, показали многообещающие результаты, хотя и немного ниже, чем их контролируемые аналоги. Их более низкие показатели производительности могут указывать на более высокий уровень ложных срабатываний, что потенциально может повлиять на их эффективность в определенных сценариях. Однако преимущество обучения без учителя заключается в его применимости в ситуациях, когда маркированные данные недостаточны или недоступны, что является распространенной проблемой на ранних стадиях обнаружения мошенничества, когда маркированных примеров недостаточно. Несмотря на общую более низкую эффективность, эти модели по-прежнему могут играть решающую роль в первоначальном обнаружении и понимании новых типов мошенничества.

Методы обучения с подкреплением, такие как модель Deep Q-Network (DQN), представляют собой многообещающий альтернативный подход. Исследование показало, что, хотя эта модель не превосходит лучшие модели обучения с учителем, она заметно превосходит методы без учителя.

Гибкость моделей обучения с подкреплением делает их особенно ценными в динамичных средах, где тактика мошенничества постоянно развивается. Способность обучения с подкреплением изучать и оптимизировать стратегии посредством взаимодействия с окружающей средой делает его универсальным инструментом в борьбе с мошенничеством с банковскими картами.

Вывод. Одним из ключевых выводов этого исследования является потенциал гибридного подхода. Объединив сильные стороны методов контролируемого, неконтролируемого обучения и обучения с подкреплением, можно создать более надежную и сложную систему обнаружения мошенничества.

Например, модели обучения с учителем можно использовать для точного обнаружения на основе помеченных данных, а модели без учителя можно применять для обнаружения аномалий при отсутствии меток. Обучение с подкреплением, в свою очередь, может динамически совершенствовать стратегии обнаружения на основе обратной связи от меняющейся среды мошенничества.

Выбор подходящего метода обнаружения мошенничества с банковскими картами зависит от различных факторов, включая доступность маркированных данных, необходимость обнаружения в режиме реального времени и способность адаптироваться к развивающимся моделям мошенничества.

Методы обучения с учителем оказываются эффективными, когда доступны высококачественные размеченные данные, в то время как подходы к обучению без учителя выгодны для выявления ранее неизвестных моделей мошенничества. Хотя обучение с подкреплением менее изучено, оно открывает возможности для разработки динамических и адаптивных стратегий обнаружения мошенничества. Исследователи и практики могут рассмотреть возможность интеграции нескольких методов, чтобы использовать свои сильные стороны и смягчить их ограничения.

Будущие исследования могут быть сосредоточены на разработке передовых гибридных или ансамблевых подходов, которые сочетают в себе различные методы, тем самым повышая точность и эффективность методов обнаружения мошенничества с банковскими картами.

Такие комплексные подходы потенциально способны решить многогранные проблемы, возникающие в результате мошеннической деятельности в финансовой сфере.

В конечном счете, выбор метода обнаружения мошенничества должен основываться на всестороннем понимании доступных данных, конкретных требований области применения и компромисса между различными методами с точки зрения производительности, адаптируемости и вычислительной сложности.

Библиографический список:

1. Бхаттачарья С., Джха С., Таракуннель К., и Westland, JC 2017. Интеллектуальный анализ данных для мошенничества с банковскими картами: сравнительное исследование. Системы поддержки принятия решений 50 (3): 602–613.
2. Болтон Р. Дж, Рука DJ, и другие. 2021. Неконтролируемые методы профилирования для обнаружения мошенничества. банковскими скоринг и кредитный контроль VII 235–255.
3. Брейман Л. 2001. Случайные леса. Машинное обучение 45(1):5–32.
4. Контролируемое и неконтролируемое обучение. [Электронный ресурс]. URL: https://www.ibm.com/think/topics/supervised-vs-unsupervised-learning
5. Чан П.К., Фан В., Продромидис А.Л. и Столфо С. Дж. 2019. Распределенный анализ данных при обнаружении мошенничества с банковскими картами. Интеллектуальные системы IEEE и их приложения 14 (6): 67–74.
6. Кортес К., Вапник В. 2015. Сети опорных векторов. Машинное обучение 20(3):273–297.
7. Даль Поццоло А., Боракки Г., Кэлен О. Алиппи К. и Бонтемпи Г. 2018. Обнаружение мошенничества с банковскими картами: реалистичное моделирование и новая стратегия обучения. Транзакции IEEE в нейронных сетях и системах обучения 29(8).
8. Обучением с подкреплением. [Электр. ресурс]. URL: https://habr.com/ru/articles/437020/
9. Дэн Л., Зельцер М.Л., Ю, Д .; Асеро А., Мохамед А.-Р. и Хинтон Г. 2020. Двоичное кодирование речевых спектрограмм с использованием глубокого автокодировщика. На одиннадцатой ежегодной конференции Международной ассоциации речевой коммуникации.
10. Камаль Б.,Фатеме Д.Автоэнкодеры и их применение в машинном обучении:обзор. Обзор искусственного интеллекта. 2024, Т. 57, 28https://link.springer.com/article/10.1007/s10462-023-10662-6
11. Дорронсоро Дж. Р., Джинель Ф. Санчес С.R., и Санта-Крус, К. 2016. Нейронное обнаружение мошенничества при операциях с банковскими картами. Транзакции IEEE в нейронных сетях.
12. Фиоре У. Де Сантис А. Перла Ф., Дзанетти П. и Палмиери Ф. 2018. Использование генеративных состязательных сетей для повышения эффективности классификации при обнаружении мошенничества с банковскими картами. Информационные науки.
13. Метод опорных векторов SVM.[Электр. ресурс].https://scikit-learn.ru/1-4-support-vector-machines/
14. Гудфеллоу И., Пуже-Абади, Дж., Мирза М., Сюй, Б. Уорд-Фарли Д., Озаир, С. Курвиль А., и Bengio, Y. 2021. Генеративные состязательные сети. В достижениях в области нейронных систем обработки информации, 2672–2680.
15. Кривко М. 2016. Гибридная модель для систем обнаружения мошенничества с пластиковыми картами. Экспертные системы с приложениями 37(8):6070–6076.
16. Мид, Адриан и Льюрис, Тайлер и др. 2018 г., Обнаружение мошенничества в состязательной среде: подход к обучению с подкреплением. Симпозиум по проектированию систем и информационной инженерии (SIEDS) 2018 года.
17. Пумсирират А., Ян Л. 2019. Обнаружение мошенничества с банковскими картами с использованием глубокого обучения на основе автоматического кодировщика и ограниченной машины Больцмана. Международный журнал передовых компьютерных наук и приложений 9 (1).
18. Цзяньцин Ф., Чжаоран В., Юйчэнь С., Чжуоран., 2020.Теоретический анализ глубокого Q-обучения. 2-й конференции по обучению динамике и управлению, PMLR 120: 486–489, 2020. URL: https://proceedings.mlr.press/v120/yang20a.html
19. https://medium.com/@cedric.vandelaer/reinforcement-learning-dqn-part-1-2-aac5f1e6e3be Обучение с подкреплением: DQN.[Электронный ресурс].
20. А.К.Бансен, Д.Ауада, А.Стоянович, Б.Оттерстен. Стратегии разработки функций для обнаружения мошенничества с кредитными картами, Expert Systems with Applications, 2016, Т.51, №1, с.134–142.

References:

1. Bhattacharya, S.; Jha, S.; Tarakunnel, K.; and Westland, J.C. Data mining for bank card fraud: A comparative study. *Decision Support Systems* 2017; 50(3):602–613.
2. Bolton, R. J.; Hand, D. J.; et al. 2021. Unsupervised profiling techniques for fraud detection. Bank Scoring and Credit Control VII 235–255.
3. Breiman, L. Random forests. *Machine Learning* 2001;45(1):5–32.
4. Supervised and Unsupervised Learning. [El resource]. https://www.ibm.com/think/topics/supervised-vs-unsupervised-learning
5. Chan, P. K.; Fan, W .; Prodromidis, A. L.; and Stolfo, S. J. Distributed data mining in bank card fraud detection. *IEEE Intelligent Systems and Applications* 2019; 14(6):67–74.
6. Cortes, C., and Vapnik, V. Support Vector Networks. *Machine Learning*. 2015; 20(3):273–297.

7. Dal Pozzolo A.; Boracchi G. Kalen O.; Alippi C.Bontempi G. Bank Card Fraud Detection:Realistic Modeling and a New Training Strategy. *IEEE Transactions on Neural Networks and Learning Systems* 2018; 29(8).
8. Reinforcement Learning. [Electronic resource]. URL: https://habr.com/ru/articles/437020/
9. Deng L.; Seltzer M.L.; Yu D.; Acero A.; Mohamed A.-r., Hinton G. 2020. Binary encoding of speech spectrograms using a deep autoencoder. In Eleventh Annual Conference of the International Speech Communication Association.
10. Kamal B., Fatemeh D. Autoencoders and their applications in machine learning: A review. *Artificial Intelligence Review*, 2024;57, 28 URL:https://link.springer.com/article/10.1007/s10462-023-10662-6
11. Dorronsoro, J. R.; Ginelle, F .; Sanchez, C. R.; and Santa Cruz, C. 2016. Neural fraud detection in bank card transactions. IEEE Transactions on Neural Networks.
12. Fiore, U.; De Santis, A .; Perla, F.; Zanetti, P.; and Palmieri, F. 2018. Using Generative Adversarial Networks to Improve Classification Performance in Bank Card Fraud Detection. Information Sciences.
13. Support Vector Machines (SVM). [El. resource]. URL: https://scikit-learn.ru/1-4-support-vector-machines/
14. Goodfellow I.; Pouget-Abadie J.; Mirza M.; Xu B. Ward-Farley, D.; Ozair, S; Courville A; Bengio, Y.2021 Generative Adversarial Networks. *In Advances in Neural Information Processing Systems*, 2672–2680.
15. Krivko, M. 2016. A Hybrid Model for Plastic Card Fraud Detection Systems. *Expert Systems with Applications* 37(8):6070–6076. (In Ross)
16. Mead Adrian and Lewis, Tyler et al. 2018. “Fraud Detection in Adversarial Environments: A Reinforcement Learning Approach.” 2018. Systems and Information Engineering Design Symposium (SIEDS).
17. Phumsirirat A.,Yang L.Bank Card Fraud Detection Using Deep Learning Based on Autoencoder and Restricted Boltzmann Machine. *International Journal of Advanced Computer Science and Applications* 2019; 9(1).
18. Jianqing, F., Zhaoran, W., Yuchen, S., Zhuoran, J., A Theoretical Analysis of Deep Q-Learning. 2nd Conference on Dynamics and Control Learning, *PMLR.* 2020;120:486–489. https://proceedings.mlr.press/v120/yang20a.html
19. Reinforcement Learning: DQN.[El. resource].https://medium.com/@cedric.vandelaer/reinforcement-learning-dqn-part-1-2-aac5f1e6e3be
20. A.K. Bahnsen, D. Aouada, A. Stojanovic, and B. Ottersten. Feature Engineering Strategies for Credit Card Fraud Detection., *Expert Systems with Applications*, 2016; 51 (1):134–142,

Сведения об авторе:
Абдурахман Джамал Джама, аспирант, кафедра «Информационная безопасность», jamaljolevas14psg@gmail.com
Information about author:
Abdurahman Jamal Jama, Graduate, Department of Information Security, jamaljolevas14psg@gmail.com
Конфликт интересов/Conflict of interest.
Автор заявляет об отсутствии конфликта интересов/The author declare no conflict of interest.
Поступила в редакцию/Received 02.09.2024.
Одобрена после/рецензирования Reviced 17.10.2024.
Принята в печать/ Accepted for publication 17.10.2024.