**Министерство науки и высшего образования**

**Российской Федерации**

**Федеральное государственное автономное образовательное учреждение высшего образования**

**«Российский университет дружбы народов имени Патриса Лумумбы»**

Факультет физико-математических и естественных наук

Кафедра информационных технологий

«Допустить к защите»

И.о. заведующего кафедрой  
информационных технологий  
д.ф.-м.н., профессор  
\_\_\_\_\_\_\_\_\_\_\_\_ Ю.Н. Орлов

« » 20 г.

**Выпускная квалификационная работа**

**бакалавра**

Направление 09.03.03 «Прикладная информатика»

ТЕМА «Искусственный интеллект как технология противодействия мошенничеству в банковской сфере»

Выполнил студент **Кипчакбаев Рамил Маратович**

|  |  |
| --- | --- |
| Группа НПИбд-01-19 | Руководитель выпускной  квалификационной работы |
| Студ. билет №1032192949 | к. ф-м. н., доц. Шорохов С. Г. |
|  |  |
|  | Автор Кипчакбаев Р.М. |

г. Москва

2023

**Федеральное государственное автономное образовательное учреждение**

**высшего образования**

**«Российский университет дружбы народов»**

**Аннотация**

**выпускной квалификационной работы**

Кипчакбаева Рамила Маратовича

на тему: Искусственный интеллект как технология противодействия мошенничеству в банковской сфере

В современном мире банковская сфера является одной из наиболее уязвимых областей в отношении мошенничества. Совершение мошеннических операций в банковской сфере часто приводит к серьезным финансовым потерям как для клиентов, так и для банков. Для борьбы с мошенничеством в банковской сфере используются различные методы и технологии, одним из которых является искусственный интеллект.

Одной из ключевых задач данной дипломной работы является исследование возможностей применения искусственного интеллекта для противодействия мошенничеству в банковской сфере. Для достижения этой цели в работе был проведен обзор существующих методов и алгоритмов, используемых для обнаружения мошеннических операций в банковской сфере. В ходе обзора были рассмотрены различные подходы к обнаружению мошенничества, включая статистические методы, машинное обучение и нейронные сети.

Объем дипломной работы составляет TODO страниц. Работа состоит из: введения, теоретической главы, главы проектирования системы, главы реализации проекта, заключения.

Во введении обосновывается актуальность темы, ставятся цели и задачи исследования, а также описываются методы исследования

В первом разделе определяются понятия мошенничества в банковской сфере, раскрываются основные способы мошенничества, анализируются последствия и риски, связанных с мошенничеством в банковской сфере, оценены преимущества и недостатки использования искусственного интеллекта в банковской сфере

Во втором разделе рассматриваются существующие методы борьбы с мошенничеством, а также современные технологии ИИ и их применение в банковском секторе.

В третьем разделе представлена реализация системы предотвращения мошеннических транзакций, оценка её эффективности и сравнение с существующими системами.

В заключении обобщены результаты исследований в данной области, сделаны выводы о применимости и эффективности искусственного интеллекта в борьбе с мошенничеством в банковской сфере и предложены рекомендации по применению ИИ для выявления мошеннических действий.

Автор ВКР

(Подпись) (ФИО)

TODO заполнить номера страниц

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ](#_Toc102888864) 3

[1. ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ](#_Toc102888865)

**[1.1. Сущность, содержание, основные виды мошенничества и его причины](#_Toc102888866)**

[**1.2. Искусственный интеллект и его применение в банковской сфере**](#_Toc102888867)

[**1.3. Проблемы и вызовы, связанные с применением искусственного интеллекта в борьбе с мошенничеством**](#_Toc102888868)

[2. АНАЛИТИЧЕСКИЙ РАЗДЕЛ](#_Toc102888869)

[**2.1. Анализ состояния мошенничеств в банковской сфере**](#_Toc102888870)

[**2.2. Анализ наиболее популярных методов обнаружения мошенничества в банковской сфере**](#_Toc102888871)

[**2.3. Анализ и оценка эффективности существующих решений на основе искусственного интеллекта для борьбы с мошенничеством в банковской сфере**](#_Toc102888871)

[3. ПРОЕКТНЫЙ РАЗДЕЛ](#_Toc102888873)

[**3.1. Описание предметной области**](#_Toc102888874)

[**3.2.** **Разработка модели искусственного интеллекта для противодействия мошенничеству в банковской сфере**](#_Toc102888874)

[**3.3. Оценка эффективности разработанной модели и сравнение ее с существующими решениями**](#_Toc102888876)

[ЗАКЛЮЧЕНИЕ](#_Toc102888877)

[Список литературных источников](#_Toc102888878)

[**Приложение 1.**](#_Toc102888879)

[**Приложение 2.**](#_Toc102888880)

# **ВВЕДЕНИЕ**

В настоящее время банковская сфера является одной из самых важных и ответственных в экономике любой страны. Однако, с развитием технологий и электронных платежных систем, появились и новые угрозы, связанные с мошенничеством и киберпреступлениями. Мошеннические операции наносят значительный ущерб не только банкам, но и клиентам, а также всей экономике в целом. Для противодействия этим угрозам, необходимо использовать современные инструменты, включая искусственный интеллект.

Цель данной дипломной работы заключается в исследовании и описании возможностей использования искусственного интеллекта для борьбы с мошенничеством в банковской сфере. В работе будет рассмотрен анализ существующих методов и технологий, применяемых в банковской сфере для обнаружения мошеннических операций. Также будет проведено исследование применения искусственного интеллекта в банковской сфере, его возможностей и преимуществ в борьбе с мошенничеством.

В работе будет проанализирована практическая реализация систем искусственного интеллекта в банковской сфере, а также рассмотрены примеры их успешного применения. Результаты исследования помогут выявить перспективы использования искусственного интеллекта в борьбе с мошенничеством и предложить пути оптимизации и совершенствования существующих методов.

В целом, дипломная работа посвящена актуальной проблеме защиты банковской сферы от мошенничества, и нацелена на исследование и разработку инновационных решений с использованием современных технологий искусственного интеллекта.

Ключевыми целями дипломной работы являются:

1. Анализ методов и технологий, используемых в банковской сфере для обнаружения мошеннических операций.
2. Исследование возможностей и преимуществ применения искусственного интеллекта в борьбе с мошенничеством в банковской сфере.
3. Анализ практической реализации систем искусственного интеллекта в банковской сфере, и рассмотрение примеров успешного применения.
4. Разработка и тестирование собственной системы искусственного интеллекта для обнаружения мошеннических операций в банковской сфере.
5. Определение перспектив применения искусственного интеллекта в борьбе с мошенничеством в банковской сфере.
6. Предложение рекомендаций по оптимизации и совершенствованию существующих методов искусственного интеллекта в борьбе с мошенничеством.

Система, которая должна предотвращать мошенничество в банковской сфере, должна отвечать ряду требований, включая:

1. Надежность: система должна быть высоконадежной и обеспечивать высокую степень защиты от кибератак и мошенничества.
2. Быстродействие: система должна быть быстрой и обрабатывать транзакции в режиме реального времени, чтобы оперативно обнаруживать мошеннические операции.
3. Гибкость: система должна быть гибкой и способной адаптироваться к изменениям в банковской сфере и новым видам мошенничества.
4. Масштабируемость: система должна быть масштабируемой и способной обрабатывать большие объемы данных, чтобы обнаруживать мошеннические операции в режиме реального времени.
5. Соответствие законодательству: система должна соответствовать законодательству и требованиям регуляторных органов в отношении защиты персональных данных и борьбы с мошенничеством в банковской сфере.

**1. ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ**

**1.1 Сущность, содержание, основные виды мошенничества, его причины и последствия**

**Мошенничество** – это редкое, хорошо продуманное, незаметно скрываемое, развивающееся во времени и часто тщательно организованное преступление, проявляющееся во многих формах[1].

Исходя из определения, можно выделить пять характеристик, вызывающие проблемы, связанными с разработкой системы обнаружения мошенничества.

Первая подчеркнутая характеристика и связанная с ней проблема касается того факта, что мошенничество встречается редко. Независимо от точной обстановки или приложения, только меньшинство рассматриваемых дел обычно касается мошенничества, из которых, кроме того, будет известно, что мошенничество касается лишь ограниченного числа. Это затрудняет как обнаружение мошенничества, поскольку мошеннические случаи покрываются не мошенническими, так и изучение исторических случаев для создания мощной системы обнаружения мошенничества, поскольку доступно лишь несколько примеров.

Второй характеристикой является тот факт, что мошенники стараются слиться с остальными и не вести себя иначе, чтобы не быть замеченными и оставаться прикрытыми обычными людьми. Это эффективно делает мошенничество незаметным, поскольку мошенникам действительно удается скрыться, хорошо обдумав и спланировав, как именно совершить мошенничество. Их поведение определенно не является импульсивным и незапланированным, иначе их было бы гораздо легче обнаружить.

Третьей характеристикой является то, что мошенники также адаптируют и совершенствуют свои методы, что им необходимо делать, чтобы оставаться незамеченными. Системы обнаружения мошенничества совершенствуются и учатся на примерах. Таким образом, методы и уловки, которые используют мошенники, развиваются со временем вместе с механизмами обнаружения мошенничества, а иногда даже опережают их. Эта игра в кошки-мышки между мошенниками и борцами с мошенничеством может показаться бесконечной игрой, но альтернативного решения пока нет. Внедряя и развивая передовые механизмы обнаружения и предотвращения мошенничества, организациям удается сокращать потери из-за мошенничества, потому что мошенники, как и другие преступники, склонны искать легкий путь и будут искать другие, более легкие возможности. Таким образом, борьба с мошенничеством путем создания передовых и мощных систем обнаружения, безусловно, не является бессмысленным усилием, но следует признать, что это, скорее всего, усилие без конца.

Четвёртой характеристикой является то, что мошенничество часто является тщательно организованным преступлением, а это означает, что мошенники часто не действуют независимо, имеют союзников и могут подстрекать подражателей. Кроме того, некоторые виды мошенничества, такие как отмывание денег и карусельное мошенничество, включают сложные структуры, созданные для организованного совершения мошенничества. Это делает мошенничество не изолированным событием, и поэтому для обнаружения мошенничества следует учитывать контекст (например, социальную сеть мошенников).

Последний элемент описания мошенничества указывает на множество различных форм, в которых происходит мошенничество. Это относится как к широкому набору методов и подходов, используемых мошенниками, так и ко многим различным условиям, в которых происходит мошенничество, или к экономической деятельности, которая подвержена мошенничеству.

В таблице 1.1 представлен неполный обзор и описание ряда типов мошенничества, которые встречаются чаще всего, или же имеют большой финансовый ущерб.

Таблица 1.1. Самые распространённые виды мошенничества

|  |  |
| --- | --- |
| Мошенничество с кредитными картами | В мошенничестве с кредитными картами происходит несанкционированное получение чужой кредитной карты. Некоторыми распространенными подвидами мошенничества с кредитными картами являются подделка кредитных карт, использование утерянных или украденных карт или получение карты мошенническим путем по почте. Можно выделить два подтипа: (1) Мошенничество с приложениями, когда люди получают новые кредитные карты от компаний-эмитентов, используя ложную личную информацию, а затем тратят как можно больше за короткий промежуток времени. (2) Поведенческое мошенничество, когда сведения о законных картах получены обманным путем, а продажи осуществляются на основе «отсутствия держателя карты». Это не обязательно требует кражи физической карты, достаточно кражи учетных данных карты. Поведенческое мошенничество касается большей части мошенничества с кредитными картами. Также встречается мошенничество с дебетовыми картами, хотя и реже. Мошенничество с кредитными картами является формой кражи личных данных, как будет определено ниже. |
| Страховое мошенничество | Данный вид мошенничества относится к любому виду страхования, как со стороны покупателя, так и продавца договора страхования. Страховое мошенничество со стороны эмитента (продавца) включает в себя продажу полисов от несуществующих компаний, непредставление премий и взбалтывание полисов для увеличения комиссионных. Мошенничество с покупателем включает в себя преувеличенные требования (страхование имущества: получение платежа на сумму, превышающую стоимость уничтоженного имущества), фальсификацию истории болезни (медицинское страхование: поддельные травмы), просроченные полисы, инсценировку смерти, похищения или убийства (мошенничество со страхованием жизни), и фиктивный ущерб (страхование автомобилей: постановочное столкновение)[2]. |
| Коррупция | Коррупция — это злоупотребление доверенной властью (по наследству, образованию, браку, выборам, назначению и т. д.) для личной выгоды. Это определение похоже на определение мошенничества, которое обсуждалось ранее, в том смысле, что целью является личная выгода. Он отличается тем, что фокусируется на злоупотреблении доверенной властью. Определение охватывает широкий спектр различных подвидов коррупции, то есть не только коррупцию со стороны политика или государственного служащего, но также, например, со стороны генерального директора или финансового директора компании, нотариуса, руководителя группы в рабочее место, администратор или сотрудник приемной комиссии в частной школе или больнице, тренер футбольной команды и т. д. [3]. |
| Мошенничество с гарантией на товар | Гарантия на продукт — это тип гарантии, которую производитель или аналогичная сторона дает в отношении состояния своего продукта, а также относится к условиям и ситуациям, при которых ремонт или замена будут производиться в случае, если продукт не работает так, как было первоначально описано. Когда продукт не может предложить описанные функции или демонстрирует отклоняющиеся характеристики или поведение, которые являются следствием производственного процесса, а не следствием неправильного использования покупателем, производитель или поставщик может потребовать компенсацию или вознаграждение. Когда условия продукта были изменены из-за использования продукта покупателем, гарантия не распространяется. Умышленное неправильное требование компенсации или вознаграждения на основе гарантии на продукт называется мошенничеством с гарантией на продукт. |
| Мошенничество в сфере здравоохранения | Мошенничество в сфере здравоохранения включает в себя подачу нечестных исков о медицинском обслуживании с целью получения прибыли. Схемы практикующих врачей включают: отдельные лица получают субсидируемые или полностью покрываемые рецептурные таблетки, которые на самом деле не нужны, а затем продают их на черном рынке с целью получения прибыли; выставление практикующими врачами счетов за услуги, которые они никогда не оказывали; подача дублирующих заявок на одну и ту же оказанную услугу; выставление счетов за непокрываемую услугу как покрываемую услугу; внесение изменений в медицинские записи и так далее. Участники могут совершить мошенничество в сфере здравоохранения, предоставив ложную информацию при подаче заявок на программы или услуги, подделывая или продавая лекарства по рецепту, одалживая или используя чужую страховую карту и т. д. |
| Отмывание денег | Процесс изъятия доходов от преступной деятельности и придания им законности. Отмывание позволяет преступникам превращать доходы, полученные незаконным путем, в, казалось бы, законные средства. Это всемирная проблема, и ежегодно в Соединенных Штатах проходит около 300 миллиардов долларов. |
| Кража личных данных | Преступление, связанное с получением личной или финансовой информации другого лица с целью присвоения имени или личности этого лица для совершения сделок или покупок. Некоторые похитители личных данных роются в мусорных баках в поисках выписок по банковским счетам и кредитным картам; другие более высокотехнологичные методы включают доступ к корпоративным базам данных для кражи списков информации о клиентах. |
| Уклонение от уплаты налогов | Уклонение от уплаты налогов — это незаконное действие или практика неуплаты причитающихся налогов. На предприятиях уклонение от уплаты налогов может иметь место в связи с подоходным налогом, налогом на занятость, налогом с продаж и акцизным налогом, а также другими федеральными, государственными и местными налогами. Примеры практик, которые считаются уклонением от уплаты налогов, включают сознательное не уведомление о доходах или занижение доходов (т. е. заявление о меньшем доходе, чем вы фактически получили из определенного источника). |

В мире существует постоянная угроза мошенничества в банковском секторе. Мошенники постоянно ищут новые способы обмана людей и банковских учреждений. Это может происходить как через традиционные каналы, такие как фишинговые атаки и кражи личных данных клиентов, так и через новые технологии, такие как криптовалюты и смарт-контракты.

Такие мошеннические действия могут привести к серьезным последствиям, например, потеря денег клиентами банка, нарушение их личной информации и финансовых данных, ущерб репутации банка и ухудшение его финансовых показателей.

Мошенничество же в банковском секторе представляет собой незаконные действия, совершаемые с целью получения финансовой выгоды или доступа к конфиденциальной информации клиентов банка. Это может включать различные формы деятельности, такие как фишинг, кражу личных данных, кражу идентификационных данных, кибератаки и отмывание денег.

Банки предпринимают различные меры для борьбы с мошенничеством, такие как использование современных технологий защиты, повышение уровня обучения персонала и клиентов, введение двухфакторной аутентификации и многое другое. Также существуют различные организации и регуляторы, которые занимаются контролем и регулированием банковской деятельности в целом и борьбой с мошенничеством в частности.

Несмотря на все усилия, мошенничество в банковском секторе остается серьезной угрозой для клиентов и финансовых учреждений, поэтому важно оставаться бдительным и принимать меры предосторожности при работе с банками и финансовыми институтами.

Мошенничество в банковском секторе может быть вызвано различными причинами, включая:

* Финансовые трудности: некоторые люди могут обращаться к мошенничеству из-за финансовых трудностей, например, из-за задолженности по кредитам, недостатка средств или нехватки денег для оплаты счетов. Они могут совершать мошенничество, чтобы получить доступ к деньгам, которых у них нет.
* Желание обогатиться: другие люди могут совершать мошенничество в банковском секторе, потому что они хотят быстро обогатиться. Они могут использовать различные методы, например, кражу личных данных, создание фиктивных компаний или участие в финансовых пирамидах.
* Недостаточная защита данных: банки могут столкнуться с мошенничеством из-за недостаточной защиты личных данных своих клиентов. Хакеры могут использовать уязвимости в системах безопасности банка, чтобы получить доступ к конфиденциальной информации, такой как номера кредитных карт, пароли и другие данные.
* Неэффективные процессы: мошенничество может происходить из-за неэффективных процессов банка, которые могут сделать его уязвимым для различных видов атак. Например, неправильное применение процедур проверки личности клиента, отсутствие контроля за операциями и другие ошибки могут привести к мошенничеству.
* Внутреннее мошенничество: наконец, мошенничество может быть вызвано внутренними факторами, такими как действия нечестивых сотрудников банка. Это может включать укрывательство операций, отклонение денег, мошенническое перевод денег и другие подобные преступления.

Мошенничество в банковской сфере может иметь серьезные последствия для банков и их клиентов. Некоторые из возможных последствий мошенничества включают в себя:

* Финансовые потери для банка и клиентов. Мошенники могут получить доступ к деньгам клиентов и/или банка, и перевести их на свои счета. Кроме того, банк может потерять деньги в результате компенсации клиентам за ущерб, нанесенный ими мошенниками.
* Ущерб репутации банка. Когда мошенничество происходит в банке, это может нанести серьезный ущерб его репутации и доверию клиентов к нему. В результате клиенты могут потерять интерес к банковским услугам, что приведет к уменьшению прибыли банка.
* Нарушение конфиденциальности данных клиентов. Мошенники могут получить доступ к личной информации клиентов, включая их имена, адреса, социальные номера и данные о банковских счетах. Это может привести к краже личности, идентификационных краж и других проблем с безопасностью.
* Потеря доверия клиентов. Если клиенты становятся жертвами мошенников в банке, они могут потерять доверие к банку и его услугам. Это может привести к тому, что клиенты уйдут к конкурентам или полностью откажутся от использования банковских услуг.
* Юридические проблемы. Если мошенничество не было выявлено вовремя, банк может столкнуться с правовыми проблемами и судебными исками со стороны клиентов. Кроме того, банк может потерять лицензию на проведение банковских операций, если не сможет обеспечить достаточный уровень защиты от мошенничества.

В целом, мошенничество в банковской сфере может привести к серьезным финансовым, репутационным и юридическим последствиям для банков и их клиентов. Поэтому очень важно для банков использовать современные технологии, в том числе искусственный интеллект, для защиты от мошенничества.

[**1.2. Искусственный интеллект и его применение в банковской сфере**](#_Toc102888867)

Искусственный интеллект (AI) — это область науки и технологии, которая занимается созданием программ и алгоритмов, которые позволяют компьютерам "думать" и "выполнять задачи", которые ранее были свойственны только человеческому интеллекту. Иными словами, искусственный интеллект относится к созданию компьютерных систем, которые способны выполнять задачи, которые требуют интеллектуальных способностей, таких как распознавание речи, обработка естественного языка, анализ данных, обучение, адаптация к новой информации и т.д.

Развитие искусственного интеллекта началось в 1950-х годах, когда были разработаны первые алгоритмы машинного обучения. Профессор Джон Маккарти из Массачусетского технологического института (MIT) создал термин "искусственный интеллект" и начал работать над созданием первых ИИ-систем. В те годы исследования в области ИИ были ограничены доступными вычислительными мощностями, и большинство задач, которые можно было решить, были простыми.

В 1960-е годы исследователи начали создавать системы, способные использовать знания и вывести логические выводы. Одной из наиболее известных систем того времени была Dendral, которая использовалась для анализа спектров молекул. В это время также была создана искусственная нейронная сеть, которая позже стала основой для различных алгоритмов машинного обучения.

В 1970-х годах исследователи начали создавать экспертные системы - ИИ-системы, которые могут принимать решения на основе знаний, которые им предоставляют эксперты в соответствующих областях. Одной из первых экспертных систем была MYCIN, разработанная в Стэнфордском университете, которая использовалась для диагностики инфекций крови.

В 1980-х годах исследователи начали активно изучать методы машинного обучения, которые позволяют компьютерным системам обучаться на основе данных. Это привело к созданию новых алгоритмов, таких как нейронные сети и метод опорных векторов, которые используются в настоящее время для обработки и анализа данных.

В 1990-х годах искусственный интеллект начал проникать в повседневную жизнь, например, в системы распознавания голоса, игры и системы управления производством. Исследователи начали создавать системы, способные использовать не только структурированные данные, но и неструктурированные данные, такие как тексты, изображения и видео. Это привело к созданию новых методов обработки естественного языка, распознавания образов и обработки изображений.

С 2000-х годов искусственный интеллект стал незаменимым в многих отраслях, таких как медицина, финансы, транспорт и реклама. Были созданы новые методы машинного обучения, такие как глубокое обучение и свёрточные нейронные сети. Эти методы позволяют обрабатывать большие объемы данных и извлекать из них более сложные зависимости. ИИ используется для оптимизации бизнес-процессов, повышения качества продуктов и услуг, анализа больших данных и создания интеллектуальных ассистентов.

В настоящее время исследования в области ИИ продолжаются, и эксперты ожидают, что ИИ станет все более широко применяемым в различных областях, включая здравоохранение, финансы, образование и др. Одним из главных вызовов, стоящих перед исследователями, является разработка более эффективных алгоритмов машинного обучения и улучшение способности ИИ систем к обучению на меньших объемах данных.

Технически искусственный интеллект может включать в себя множество различных методов и технологий, таких как машинное обучение, глубокое обучение, нейронные сети, алгоритмы генетической оптимизации, различные формы логики, робототехника и многие другие. Все эти методы и технологии предназначены для того, чтобы обучать компьютерные системы находить закономерности в данных, выделять наиболее значимые признаки и принимать решения на основе полученных знаний.

В основе искусственного интеллекта лежит машинное обучение (Machine Learning), которое позволяет компьютерам обучаться на больших объемах данных, выявлять образцы и закономерности и делать предсказания на основе этих данных. Машинное обучение включает в себя алгоритмы классификации, регрессии, кластеризации, нейронные сети, глубокое обучение и т.д.

Искусственный интеллект находит широкое применение в различных сферах, включая:

* Медицина. Искусственный интеллект может использоваться для анализа медицинских изображений, диагностики заболеваний и разработки новых лекарственных препаратов. Он также может помочь в улучшении точности диагностики и прогнозировании исхода лечения.
* Финансы. Искусственный интеллект может использоваться для анализа и прогнозирования экономических трендов, определения рисков и принятия решений в области инвестирования.
* Транспорт. Искусственный интеллект может быть использован для создания автомобилей, способных к самостоятельному управлению, что повышает безопасность дорожного движения. Также он может быть использован для оптимизации маршрутов, управления транспортными потоками и улучшения общественного транспорта.
* Реклама и маркетинг. Искусственный интеллект может использоваться для анализа поведения потребителей и прогнозирования их предпочтений, что позволяет создавать более точные рекламные и маркетинговые кампании.
* Производство. Искусственный интеллект может использоваться для оптимизации производственных процессов, улучшения качества продукции, прогнозирования неисправностей оборудования и улучшения производственной безопасности.
* Образование. Искусственный интеллект может быть использован для создания персонализированных образовательных программ, анализа учебных материалов и повышения эффективности обучения.
* Автомобильная промышленность: искусственный интеллект может помочь в разработке более безопасных и удобных автомобилей, а также в создании самоуправляемых автомобилей.
* Энергетика: искусственный интеллект может помочь в оптимизации расходов энергии и увеличении эффективности энергетических систем.
* Космическая индустрия: искусственный интеллект может помочь в управлении миссиями в космосе, анализе данных и разработке новых технологий.
* Безопасность: искусственный интеллект может помочь в обнаружении и предотвращении кибератак, а также в обеспечении безопасности на транспорте и в общественных местах.

Искусственный интеллект имеет огромный потенциал для предотвращения мошенничества в банковской сфере. Ниже перечислены несколько преимуществ использования ИИ для предотвращения мошенничества в банковском секторе:

* Важным преимуществом использования искусственного интеллекта в предотвращении мошенничества является автоматизация процесса. Традиционно, выявление мошенничества в банковской сфере требует большого количества времени и усилий со стороны специалистов, которые должны вручную анализировать большие объемы данных и выявлять необычные паттерны и аномалии. Использование ИИ позволяет автоматизировать этот процесс и значительно ускорить его. Системы ИИ могут самостоятельно обрабатывать и анализировать данные, что сокращает необходимость вручную проверять каждую транзакцию. Например, системы машинного обучения могут обучаться на основе большого количества данных о транзакциях, которые содержат как мошеннические, так и не мошеннические операции. Это позволяет им создавать модели, которые автоматически выявляют потенциальные случаи мошенничества и оповещают об этом соответствующих сотрудников банка. Кроме того, автоматизация процесса позволяет банкам снизить затраты на обучение и найм квалифицированных специалистов, позволяя сократить время и затраты на анализ данных и повысить эффективность борьбы с мошенничеством. Системы ИИ работают круглосуточно без прерываний, что обеспечивает более быструю реакцию на возможные случаи мошенничества.
* Ещё одним из ключевых преимуществ использования искусственного интеллекта является быстрое реагирование на мошеннические операции. Системы искусственного интеллекта могут быстро анализировать большие объемы данных, что позволяет выявлять аномалии и необычные транзакции, которые могут быть связаны с мошенничеством. Например, ИИ может автоматически анализировать все транзакции, совершенные клиентом, и выявлять необычные паттерны поведения, такие как неожиданные покупки или снятие денег с разных мест в разное время. Быстрое реагирование на мошеннические операции позволяет банкам принимать меры для предотвращения ущерба, связанного с мошенничеством, в режиме реального времени. Например, банк может заблокировать счет клиента или отменить транзакцию, которая выглядит подозрительно, еще до того, как мошенник успеет получить доступ к деньгам.
* Следующим преимуществом использования искусственного интеллекта в предотвращении мошенничества является адаптивность систем ИИ к новым видам мошенничества. Мошенники постоянно разрабатывают новые методы атаки на банки и их клиентов. Это может быть использование новых технологий, новых видов социальной инженерии или других способов обмана. В связи с этим, традиционные методы борьбы с мошенничеством могут стать устаревшими и перестать быть эффективными. Системы ИИ могут обучаться на основе данных о новых мошеннических схемах и изменять свои алгоритмы для более эффективной их обнаружения. Например, если банк обнаруживает новый вид мошенничества, система ИИ может быстро приспособиться к нему и начать обнаруживать такие операции на основе новых данных. Кроме того, системы ИИ могут использоваться для выявления необычных паттернов и аномалий, которые могут свидетельствовать о новых видах мошенничества, что обеспечивает более адаптивный и эффективный подход к обнаружению и предотвращению мошеннических операций, что повышает уровень безопасности банков и защищает интересы их клиентов. Сокращение затрат на предотвращение мошенничества. Использование ИИ может помочь банкам сократить затраты на предотвращение мошенничества, так как автоматизированные процессы могут быть более эффективными и точными, чем процессы, осуществляемые людьми.
* Защита конфиденциальности данных. Еще одним важным преимуществом использования искусственного интеллекта в предотвращении мошенничества в банковском секторе является защита конфиденциальности данных клиентов. В современном мире, когда количество персональных данных, хранимых банками, постоянно растет, защита конфиденциальности становится важнее, чем когда-либо. Как правило, данные клиентов банков содержат очень чувствительную информацию, такую как имена, адреса, номера социального страхования, номера счетов и другую конфиденциальную информацию. Системы ИИ могут использоваться для обнаружения аномалий и необычных паттернов в данных, которые могут свидетельствовать о попытках несанкционированного доступа или мошенничества. Использование ИИ в процессе мониторинга позволяет выявлять и предотвращать утечки конфиденциальной информации и защищать ее от несанкционированного доступа. Кроме того, системы ИИ могут использоваться для защиты от кибератак, таких как вирусы, троянские программы и фишинг. Они могут анализировать трафик и обнаруживать аномальные попытки доступа к системам банка, а также распознавать и блокировать вредоносные программы и угрозы, которые могут привести к утечке конфиденциальной информации.
* Улучшение качества обслуживания клиентов. Системы ИИ могут использоваться для автоматизации процесса обработки заявок на кредит, проверки идентификации клиента, определения кредитного скоринга и принятия решения о кредите. Это упрощает и ускоряет процесс принятия решения и позволяет клиентам получать более быстрый и эффективный сервис. Кроме того, системы ИИ могут использоваться для улучшения процесса онлайн-банкинга и мобильных приложений. Это может включать в себя автоматизацию процессов, таких как оплата счетов, переводы средств и мониторинг баланса счета. Благодаря использованию ИИ, клиенты могут получать персонализированные предложения и рекомендации, основанные на их истории использования банковских услуг и предпочтениях. Таким образом, использование систем ИИ позволяет банкам предоставлять более высокое качество обслуживания клиентов, ускорять процессы и повышать эффективность банковских услуг. Это может привести к улучшению удовлетворенности клиентов и повышению лояльности к банку.
* Снижение затрат. Традиционные методы борьбы с мошенничеством могут быть очень затратными, требуя большого количества ресурсов и времени. Например, вручную проверять каждую транзакцию на предмет мошенничества может быть очень дорогостоящим и требует большого количества персонала. С использованием искусственного интеллекта для предотвращения мошенничества в банковском секторе можно снизить затраты на персонал и ресурсы. Автоматизация процесса позволяет быстро обрабатывать большие объемы данных и выявлять потенциальные случаи мошенничества, что может снизить затраты на ручную обработку и анализ данных. Кроме того, применение искусственного интеллекта может снизить количество ложных срабатываний и ошибок, что позволяет сократить затраты на исправление ошибок. В итоге, использование искусственного интеллекта для предотвращения мошенничества в банковском секторе может быть выгодным для организации, позволяя сократить затраты на персонал и обработку данных, а также повысить эффективность и точность работы.

[**1.3. Проблемы и вызовы, связанные с применением искусственного интеллекта в борьбе с мошенничеством**](#_Toc102888868)

Хотя применение искусственного интеллекта в борьбе с мошенничеством может привести к значительному улучшению безопасности, но у него также есть свои проблемы и вызовы, которые могут ограничить его эффективность. Некоторые из таких вызовов включают в себя:

* Недостаток качественных данных: для того, чтобы ИИ был эффективным в борьбе с мошенничеством, необходимы качественные данные. Однако, в борьбе с мошенничеством зачастую не всегда есть доступ к достаточно широкому диапазону данных, так как мошенники постоянно изменяют свои тактики и методы.
* Сложность обучения модели: Обучение ИИ модели может быть сложным процессом, особенно если речь идет о более сложных задачах. Необходимо иметь квалифицированных специалистов, чтобы правильно настроить модели и обучить их на подходящих данных.
* Ложноположительные и ложноотрицательные результаты: Использование ИИ может привести к ложноположительным и ложноотрицательным результатам. Например, ИИ может считать подозрительной легитимную транзакцию, которая попадает в категорию аномалии. Или, наоборот, ИИ может пропустить мошенническую транзакцию, которая считается нормальной, так как мошенник может использовать новые методы и тактики. Это может приводить к недовольству клиентов и потере доверия к банковской организации.
* Необходимость постоянного обновления модели: Мошенники постоянно развивают свои методы, что требует от организаций постоянно обновлять свои ИИ модели, чтобы оставаться эффективными в борьбе с новыми мошенническими схемами.
* Проблемы безопасности: использование ИИ может также повысить риски безопасности, так как мошенники могут использовать атаки на ИИ для искажения результатов и обойти систему. Кроме того, необходимо обеспечить безопасность данных и предотвратить утечки информации, которые могут использоваться мошенниками.
* Этические и юридические вопросы: Использование ИИ в борьбе с мошенничеством также может вызвать этические и юридические вопросы. Например, сбор и обработка данных могут нарушать приватность клиентов. Кроме того, возможно, что ИИ может принимать решения, которые могут оказаться дискриминационными.

В целом, применение ИИ в борьбе с мошенничеством имеет множество преимуществ, однако, необходимо учитывать и эти вызовы и проблемы, чтобы эффективно использовать технологии ИИ в банковской сфере.

**2. АНАЛИТИЧЕСКИЙ РАЗДЕЛ**

**2.1 Анализ состояния мошенничеств в банковской сфере**

Банковская сфера является одной из наиболее динамичных и быстро развивающихся отраслей экономики, но в то же время является наиболее уязвимой для преступных посягательств. Это связано с недостаточностью действенных механизмов правового регулирования, сокрытием реальной уязвимости банковской сферы от общества, низкой финансовой грамотностью некоторых групп населения, быстрой сменой форм обслуживания, особенно в связи с пандемией. Поскольку ущерб от мошеннических действий в банковской сфере оценивается в триллионы рублей, а количество потерпевших растет быстрыми темпами, как государство, так и общество заинтересованы в защите этого важного экономического сектора[4].

Известно, что основными преступлениями, которые происходят в банковской сфере, являются кражи. В последнее время структура этих преступлений значительно изменилась, что связано с изменением форм работы банков и организации платежных систем. С переходом от наличных денег и финансовых документов к безналичным платежам изменилась структура преступлений, совершаемых в этой сфере. Сегодня более 3,8 миллиардов человек по всему миру, что составляет 69% населения планеты, имеют счета в банках или используют услуги мобильных платежных систем. Согласно отчету Global Findex, количество людей, использующих финансовые услуги, продолжает расти, а с 2011 по 2017 год количество открытых новых счетов в банках достигло 1,2 миллиарда человек. В России за последние 10 лет количество банковских счетов увеличилось почти вдвое: с 539,4 миллиона в 2011 году до 922,3 миллиона на начало 2021 года. Почти все эти счета (99%) принадлежат физическим лицам. Объем операций по карточным счетам составляет около 100 трлн рублей[5].

В современных условиях все более распространено дистанционное банковское обслуживание, а также хранение денежных средств на банковских счетах или в виртуальном пространстве. Это открывает возможности для преступников активно использовать различные способы дистанционного доступа к этим средствам и их хищения. Из-за этого мошенничество в банковской сфере, особенно связанное с использованием электронных средств платежа, является наиболее динамично развивающимся видом хищения денежных средств. Недавно опубликованная статистика Центрального банка Российской Федерации показала, что только за январь-март 2021 года мошенники совершили несанкционированные переводы и украли у граждан и компаний в России 2,9 млрд рублей, что на 57% больше, чем за первый квартал 2020 года[6].

Некоторые недавние цифры дают представление о предполагаемых размерах и финансовых последствиях мошенничества:

* Типичная организация ежегодно теряет 5 процентов своих доходов из-за мошенничества.
* Общая стоимость страхового мошенничества (немедицинского страхования) в Соединенных Штатах оценивается более чем в 40 миллиардов долларов в год.
* Мошенничество обходится Соединенному Королевству в 73 миллиарда фунтов стерлингов в год по данным Национального управления по борьбе с мошенничеством Британии.
* Компании, выпускающие кредитные карты, «теряют примерно семь центов на каждые сто долларов транзакций из-за мошенничества» - Эндрю Шраге, Money Crashers Personal Finance, 2012.
* Средний размер неформальной экономики в процентах от официального ВНД в 2000 году в развивающихся странах составлял 41 процент, в странах с переходной экономикой — 38 процентов, а в странах ОЭСР — 18 процентов[7].

Анализ динамики исследуемых мошенничеств в РФ за последние пять лет (2016-2020 гг.) показывает устойчивый рост уровня мошенничества. За этот период абсолютное количество мошеннических преступлений увеличилось в 3,7 раза: от 8 713 преступлений, зарегистрированных в 2016 году, до 32 186 в 2020 году. Это же подтверждают и относительные показатели. Коэффициент интенсивности мошенничества за рассматриваемый период увеличился на такую же величину (в 3,7 раза), с 5,9 преступлений, приходящихся на 100 тыс. человек в 2016 году, до 21,9 в 2020 году (см. таблицу 2.1).

Однако стоит отметить, что мошенничества данного типа не распространены равномерно по всей территории страны. За период с 2016 по 2019 годы большинство мошеннических деяний в банковской сфере было зарегистрировано в Приволжском и Центральном федеральных округах (ПФО, ЦФО), что объяснимо, так как эти регионы имеют наибольшую плотность населения и наиболее развитую сеть банковских учреждений. В то же время, в федеральных округах с меньшей численностью населения и большими не заселенными территориями количество зарегистрированных мошеннических деяний было значительно меньше. Например, в Сибирском федеральном округе (СФО) было зарегистрировано на 1,7 раза меньше мошеннических деяний, чем в ПФО, а в Уральском федеральном округе (УФО) и Дальневосточном федеральном округе (ДФО) - соответственно в 3 и фактически в 6 раз меньше (см. табл. 2.2).

Таблица 2.1 Динамика состояния зарегистрированных случаев мошенничества в банковской сфере по федеральным округам за 2016–2019 гг.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Год | ЦФО | СЗФО | ЮФО | СКФО | ПФО | УФО | СФО | ДФО | РФ |
| 2016 | 2 589 | 572 | 846 | 524 | 1 846 | 788 | 1 353 | 156 | 8 713 |
| 2017 | 2 480 | 668 | 1 177 | 525 | 2 160 | 912 | 1 230 | 290 | 9 480 |
| 2018 | 2 883 | 1 250 | 1 730 | 889 | 3 439 | 1 089 | 1 974 | 479 | 13 33 |
| 2019 | 5 407 | 3 192 | 2 747 | 1 435 | 7 765 | 2 093 | 4 545 | 1 640 | 28904 |

Таблица 2.2 Динамика коэффициента интенсивности мошенничества в банковской сфере в Российской Федерации, Сибирском, Уральском, Дальневосточном федеральных округах в 2016–2020 гг. (на 100 тыс. человек)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Год | РФ | СФО | УФО | ДФО |
| 2016 | 5,9 | 7 | 6,4 | 2,5 |
| 2017 | 6,5 | 6,4 | 7,4 | 4,7 |
| 2018 | 9,3 | 10,2 | 8,8 | 3,9 |
| 2019 | 19,7 | 26,5 | 16,9 | 20 |

В 2022 году количество ресурсов, к которым был ограничен доступ на основании информации от Банка России, увеличилось в более чем три раза по сравнению с 2021 годом и достигло отметки в 10 716 единиц (рисунок 2.1)[4].

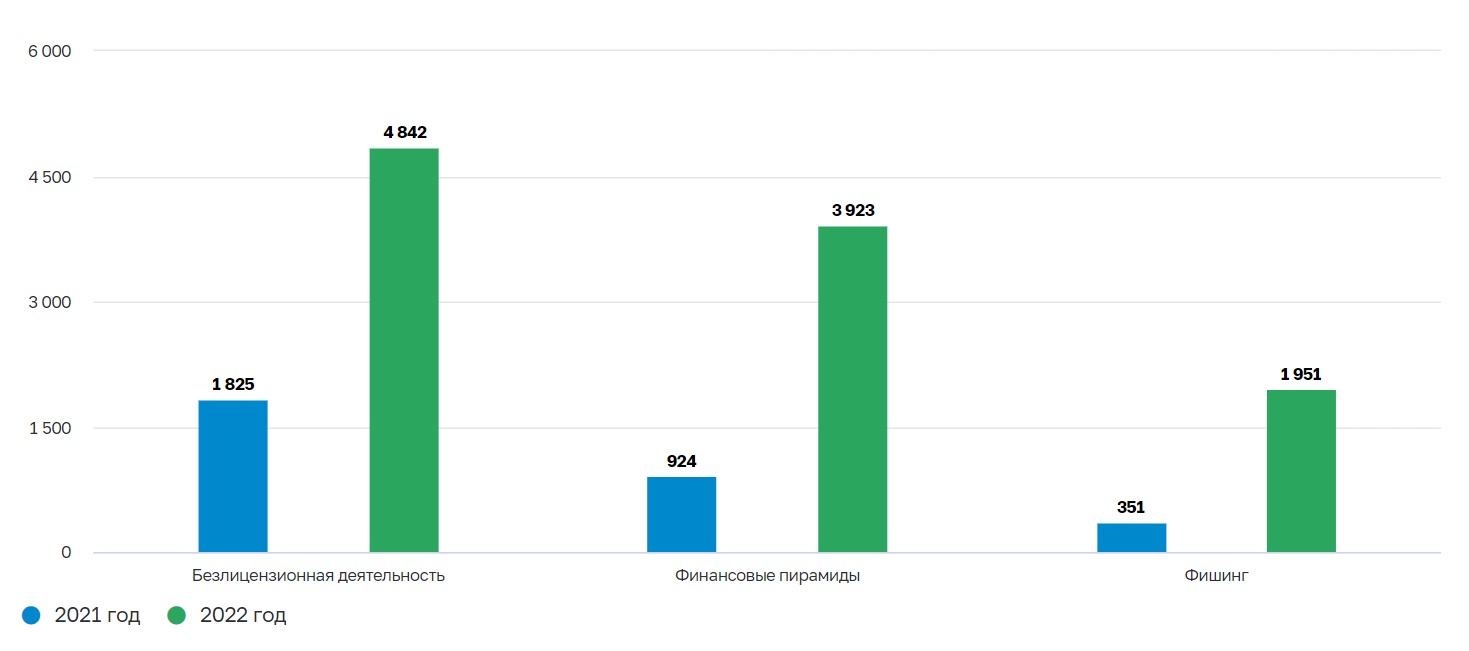


Рисунок 2.1 Мошеннические ресурсы, направленные в Генеральную прокуратуру Российской Федерации (ед.)

В 2022 году, как и в предыдущем году, основная часть ресурсов (34%) была заблокирована Банком России из-за их использования злоумышленниками для незаконной деятельности на рынке ценных бумаг и для рекламы несуществующих кредитных, микрофинансовых и страховых организаций. Примерно 27% заблокированных ресурсов относились к сайтам, которые пропагандируют финансовые пирамиды или попадали в категорию мошенничества, такие как сайты, предлагающие возможность получения государственных компенсаций или денежных выплат за прохождение онлайн-тестов и опросов, а также сайты, продававшие билеты на различные мероприятия и туры. Примерно 12% заблокированных ресурсов относились к категории "фишинг", которые злоумышленники маскировали под официальные сайты финансовых организаций. Менее 1% заблокированных ресурсов относилось к сайтам, которые распространяли вредоносное программное обеспечение (ВПО) (рисунок 2.2)[4].

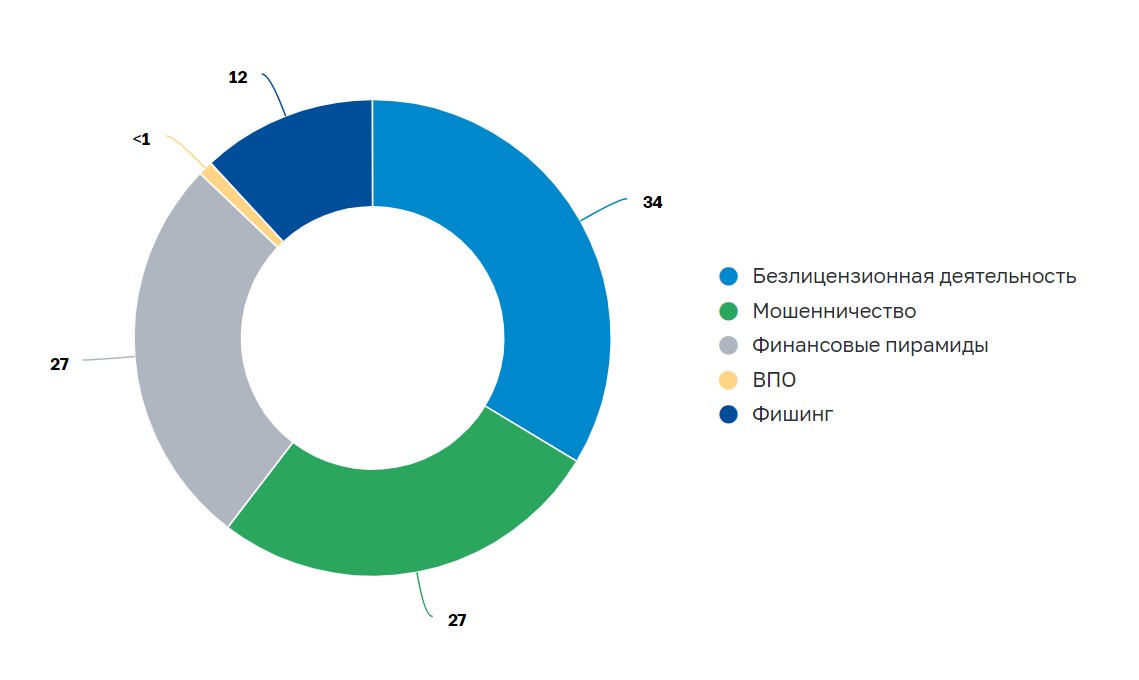


Рисунок 2.2 Типы ресурсов, используемые злоумышленниками в 2022 году (%)

Эта тенденция объясняется изменением методов, используемых мошенниками, чтобы совершать преступления, в течение исследуемого периода. Мошенники все чаще используют более недорогие психологические методы, чтобы обмануть свои жертвы. Исследование, проведенное аналитиками банка Тинькофф в 2020-2021 гг. в российской банковской сфере, выявило наиболее распространенные методы мошенничества, а также описало типичные схемы обмана. Исследование показало, что доля мошенничеств, использующих социальную инженерию, выросла более чем в 2 раза за 3 года: с 36% в 2018 году до 80% в 2020 году. В 2020 году 50% всех денег, полученных мошенниками, были переведены ими сами (в 2019 году эта доля составляла только 30%), а 28% - потери от того, что клиенты раскрывали данные карт или секретные коды. При анализе сценариев мошенничества с использованием электронных средств платежа было выявлено, что в 42% случаев мошенники представлялись службой безопасности сторонних банков, в 25% случаев - службой безопасности банка, по карте которого впоследствии будут запрашивать информацию, 11% - на схемы с покупкой и продажей товаров в сети, и 22% - на другие сценарии (инвестиции, комиссии за получение выигрыша, вознаграждения за опрос и т. д.). Средняя сумма, которую мошенники украли у одного клиента, варьировала от 48,8 тыс. руб. до 4,7 тыс. руб. в 2019 году и от 45 тыс. руб. до 4,5 тыс. руб. в 2020 году, в зависимости от способа банковского мошенничества.

В целом, можно сделать вывод, что мошенничество в банковской сфере имеет очень неблагоприятную динамику, и среди других видов преступлений, это преступление развивается наиболее активно. В ближайшие годы уровень мошенничества в банковской сфере в целом, а также его доля в общей структуре мошенничества будет продолжать расти. Этот вид мошенничества в настоящее время и в будущем будет доминирующим на территории европейской части России, связано это с наиболее развитой банковской сетью и большим количеством населения, проживающего на данной территории. Однако, в сибирских и дальневосточных регионах прирост рассматриваемых преступлений будет выше, поскольку использование банковских технологий расширит возможности для совершения преступлений на этих малонаселенных территориях, особенно в периоды вспышек пандемии и локдаунов. Более того, создание и развитие цифровых банков и электронных платежных услуг, а также возможность клиентов осуществлять самостоятельные переводы и платежи, приводит к увеличению количества транзакций в банках, и смещает интерес мошенников от банкоматов и организаций торговли в сторону CNP-транзакций и дистанционного банковского обслуживания.

**Анализ наиболее популярных методов обнаружения мошенничества в банковской сфере**

Методы обнаружения мошенничества в банковской сфере необходимы для защиты банковских клиентов и банков от потенциальных угроз, связанных с мошенничеством и кражей денег. Поэтому, методы обнаружения мошенничества в банковской сфере являются неотъемлемой частью операционной деятельности банков. Они позволяют банкам быстро и эффективно обнаруживать любую подозрительную активность, связанную с мошенничеством, и предпринимать необходимые меры для ее предотвращения. В связи с этим было разработано множество способов выявления мошенников, рассмотрим самые популярные из них:

**Экспертная система правил** (Rule-Based Expert System) - это метод обнаружения мошенничества, который использует заранее определенные правила и условия для обнаружения подозрительных транзакций или активности на счетах клиентов. Эта система работает на основе набора правил, которые могут быть созданы как людьми, так и автоматически на основе данных. Примеры подобных правил приведены ниже:

* Если у клиента есть автокредит и автострахование, то у клиента есть расчетный счет в 80 процентах случаев.
* Если клиент покупает спагетти, то в 70% случаев покупатель покупает красное вино.
* Если клиент посещает веб-страницу А, то он посетит веб-страницу Б в 90% случаев.

Экспертные системы обычно состоят из трех основных компонентов: базы знаний, машины вывода и интерфейса пользователя. В базе знаний хранится информация о предметной области, которая доступна в настоящий момент. Машина вывода отвечает за применение этих знаний к решению конкретных задач. Интерфейс пользователя обеспечивает коммуникацию между экспертной системой и пользователем, позволяя последнему задавать вопросы и получать ответы. Более подробно модель системы изображена на рисунке 2.3 В целом, экспертная система является моделью знаний эксперта в определенной области и его умения применять эту информацию для решения задач.

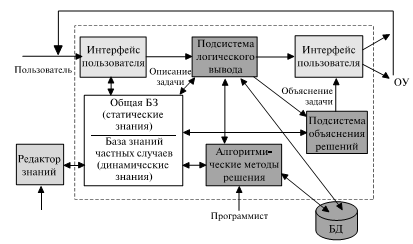


Рисунок 2.3. Структура экспертной системы

В базе знаний содержатся известные факты, представленные в виде объектов с их значениями и условиями (правилами). Эта база также может включать выражения неопределенности, такие как коэффициенты доверия, которые ограничивают достоверность каждого факта. В отличие от традиционных баз данных, база знаний имеет символьное содержание и использует логические правила.

В экспертной системе главным компонентом является механизм машины вывода, который ищет решения, основываясь на правилах рациональной логики. Этот механизм активируется пользовательским запросом и выполняет несколько задач:

**WHILE** работа не закончена

**Разрешение конфликтов**. Если имеются активизированные правила, то выбрать правило с наивысшим приоритето; В противном случае работа закончена.

**Действие**. Последовательно осуществить действия, указанные в правой части выбранного активизированного правила. В данном цикле проявляется непосредственное влияние тех действий, которые изменяют содержимое рабочей памяти. Удалить из рабочего список правил только что запущенное правило.

**Согласование**. Обновить рабочий список правил путем проверки того, выполняется ли левая часть каких-либо правил. В случае положительного ответа активизировать соответсвующие правила. Удалить активизированные правила, если левая часть соответсвующих правил больше не выполняется.

**Проверка условий останова**. Если осуществлено дейсвие halt или дана команда break, то работа закончена.

**END-WHILE**

Принять новую команду пользователя

В частности, он сравнивает информацию в запросе пользователя с информацией в базе знаний, ищет цели и причинно-следственные связи, и оценивает относительную достоверность фактов на основе связанных с ними коэффициентов доверия. Полный цикл интерпретатора изображён на рисунке 2.4.

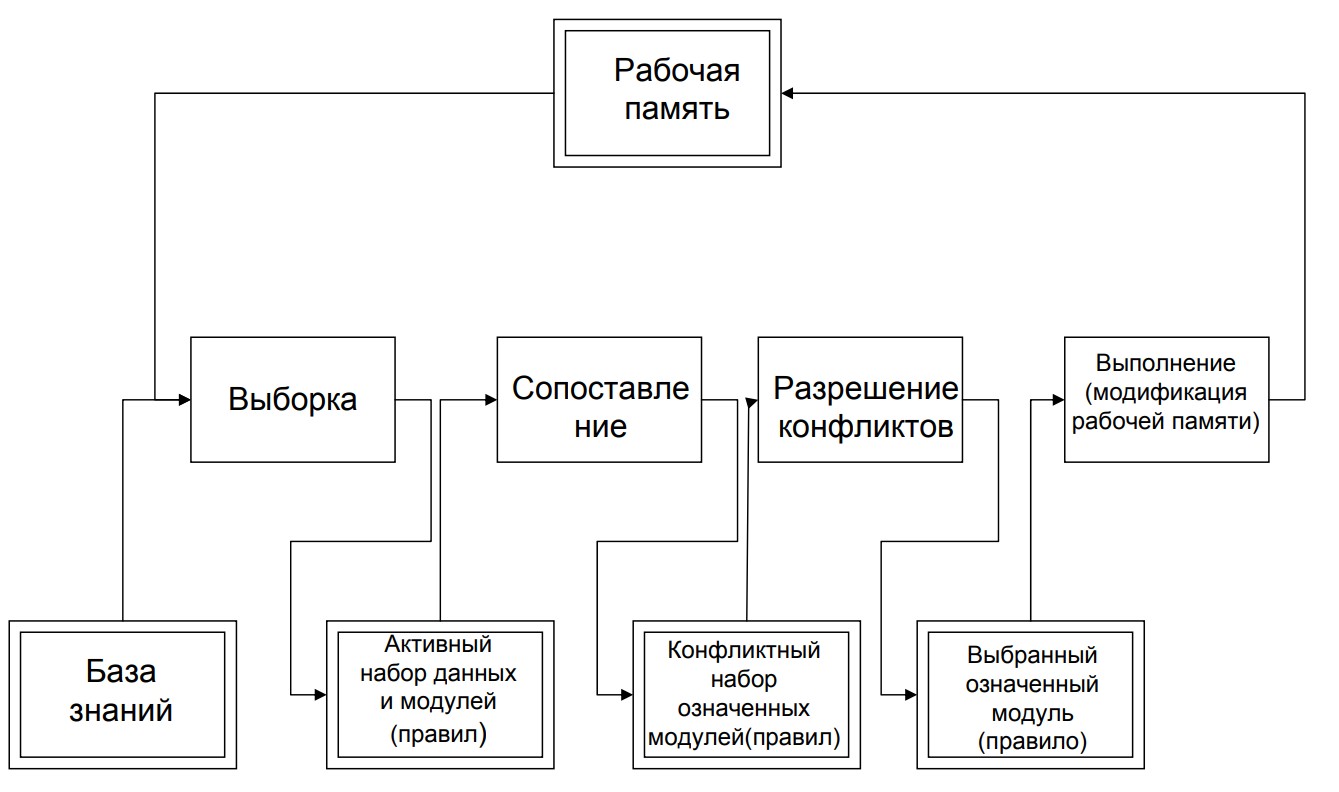


Рисунок 2.4. Цикл работы интерпретатора.

В работе ЭС можно выделить два основных режима: режим приобретения знаний и режим решения задачи (режим консультации или режим использования).

1. В процессе приобретения знаний для экспертной системы, эксперт взаимодействует с системой при помощи инженера знаний. Через компонент приобретения знаний эксперт описывает проблемную область, представляя ее в виде набора фактов и правил. То есть он заполняет систему знаниями, которые позволяют ей самостоятельно решать задачи в данной области. Стоит отметить, что в традиционном подходе к программированию этот процесс соответствует этапам алгоритмизации, программирования и отладки, которые выполняет программист. Однако, при разработке экспертной системы, задачу наполнения системы знаниями выполняет эксперт, который не обязательно обладает навыками программирования.
2. В режиме консультаций, пользователь общается с ЭС с целью получения результата и способа его получения. В зависимости от назначения ЭС, пользователь может не обладать достаточными знаниями в данной предметной области и обращаться к системе за результатом, который сам не может получить. В другом случае пользователь может быть специалистом в данной области и обращаться к ЭС с целью ускорения получения результата и избавления от рутины.

Важным этапом разработки экспертной системы является проверка корректности базы знаний – процесс, который позволяет убедиться в отсутствии внутренних ошибок, таких как избыточность, противоречивость и неполнота. Главная цель проверки продукционной БЗ заключается в выявлении аномалий, таких как противоречивых или избыточных правил, циклов и т.п. После анализа информации, полученной в результате проверки, необходимые изменения вносятся в базу для обеспечения корректного вывода. Одним из ключевых элементов проверки продукционной БЗ является выявление ошибок в цепочках вывода.

Экспертная система правил имеет несколько преимуществ по сравнению с традиционными программами:

1. Быстрота и эффективность: ЭС правил способна быстро и эффективно решать задачи в специализированной области знаний благодаря заранее определенным правилам и знаниям.
2. Точность и надежность: ЭС правил обладает высокой точностью и надежностью при решении задач благодаря использованию точных знаний и логических правил.
3. Универсальность: ЭС правил может использоваться в различных областях знаний, что делает ее универсальным решением для решения задач в разных областях.
4. Снижение затрат: ЭС правил позволяет сократить затраты на обучение и поддержку персонала, т.к. не требуется наличие специалистов в области программирования.
5. Автоматизация рутинных задач: ЭС правил позволяет автоматизировать рутинные задачи, что позволяет сократить время и улучшить эффективность работы.
6. Легкость использования: ЭС правил легко использовать, т.к. взаимодействие с ней происходит через простой интерфейс, что делает ее доступной для широкого круга пользователей.

Среди недостатков экспертных систем, которые могут снизить их эффективность и применимость, можно выделить следующие:

1. Ограниченность области применения: экспертные системы правил применяются только в тех областях, где можно сформулировать явные правила. В областях, где правила не могут быть явно сформулированы или изменяются слишком быстро, экспертные системы могут быть бесполезными.
2. Недостаточная гибкость: правила, используемые в экспертных системах, обычно являются жесткими и не учитывают контекст или изменения в предметной области. Это может ограничивать применимость системы в динамическом окружении.
3. Проблемы с нечеткой логикой: экспертные системы не всегда могут обрабатывать нечеткую логику или неопределенность в данных, что может привести к неверным выводам или неполным решениям.
4. Трудность в создании и поддержке: создание экспертных систем может быть дорогостоящим и трудоемким процессом, который требует большого количества времени и ресурсов. Кроме того, поддержка и обновление системы может также требовать значительных усилий.
5. Проблемы с обучением: для создания экспертной системы требуется экспертное знание в предметной области, и обучение системы может быть сложным процессом. Кроме того, экспертная система может стать устаревшей со временем и потребовать дополнительного обучения.

**Кластеризация** может быть полезным методом для обнаружения мошенничества в банковской сфере, поскольку она может помочь выявить группы транзакций, которые отличаются от обычных и могут быть связаны с мошенничеством.

Например, можно использовать кластеризацию для группировки клиентов по таким характеристикам, как местоположение, тип транзакции, сумма транзакции и другим факторам. Затем можно проанализировать каждый кластер на предмет аномалий и необычных транзакций. Очень популярными наборами характеристик транзакций, используемых для кластеризации при обнаружении мошенничества, являются не давность (recency), частотные (frequency) и денежные (monetary) (RFM) переменные.

При использовании для обнаружения мошенничества возможной целью кластеризации может быть группирование аномалий в небольшие разреженные кластеры. Затем их можно дополнительно проанализировать и проверить с точки зрения их характеристик и потенциально мошеннического поведения (рисунок 2.5).

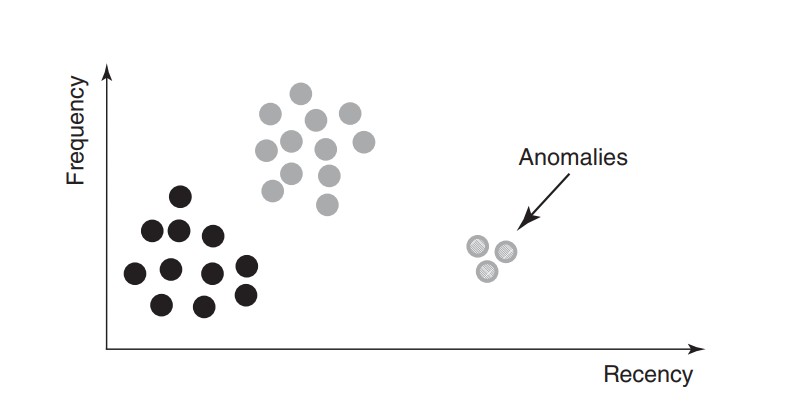


Рисунок 2.5. Кластерный анализ для обнаружения мошенничества

Как уже говорилось, цель кластеризации состоит в том, чтобы сгруппировать наблюдения на основе сходства. Следовательно, метрика расстояния необходима для количественной оценки сходства. Различные метрики расстояния были введены в литературе как для непрерывных, так и для категориальных данных.

Для непрерывных данных зачастую используются расстояние Евклида (формула 2.1) или Манхэттенское (формула 2.2), оба расстояния продемонстрированы на рисунке 2.6:

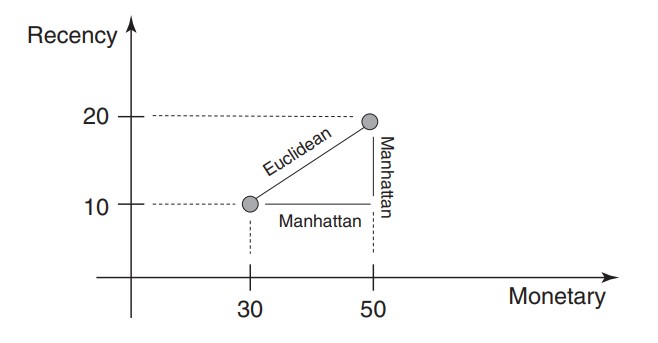


Рисунок 2.6. Евклидово и Манхэттенское расстояния.

Помимо непрерывных переменных, для кластеризации могут использоваться и категориальные переменные. Давайте сначала обсудим случай бинарных переменных. Они часто используются в методах обнаружения страхового мошенничества, которые обычно основаны на ряде предупреждающих признаков, помечающих претензию как подозрительную или нет. Предположим, у нас есть следующий набор данных с двоичными индикаторами, представленный в таблице 2.1.

Таблица 2.1. Пример категориальных переменных.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Плохая история вождения | Оплачено наличными | Доступна информация о покупке автомобиля | Покрытие увеличено | Автомобиль никогда не проверялся |
| Пункт 1 | Да | Нет | Да | Да | Нет |
| Пункт 2 | Да | Да | Нет | Нет | Нет |

Первый способ вычислить расстояние или сходство между пунктами 1 и 2 состоит в использовании простого коэффициента соответствия (Simple Matching Coefficient - SMC), который просто вычисляет количество идентичных совпадений между значениями переменных следующим образом:

В нашем примере SMC(Пункт 1, Пункт 2) = 2/5.

Однако, стоит учитывать, что кластеризация может не всегда быть эффективной для обнаружения мошенничества, особенно если мошенники умело маскируют свои действия. Кроме того, кластеризация может давать ложные срабатывания, то есть ошибочно относить обычные транзакции к кластеру мошеннических транзакций.

Таким образом, кластеризация может быть полезным инструментом для обнаружения мошенничества в банковской сфере, но ее результаты должны быть подтверждены другими методами, такими как машинное обучение и статистический анализ, для более точного выявления мошеннических транзакций.

**Деревья решений** - метод машинного обучения, который используется для прогнозирования значений целевой переменной на основе набора входных данных. Он представляет собой структуру дерева, в которой каждый узел представляет собой тест на определенное свойство, каждая ветвь соответствует возможному результату теста, а каждый лист дерева представляет собой прогноз для целевой переменной.

Дерево решений строится по обучающей выборке, где каждый элемент выборки представляет собой вектор признаков и значение целевой переменной. Алгоритм ищет лучший тест на основе выбранных критериев разделения (например, энтропия, информационный коэффициент Джини и другие), который разделяет элементы выборки на две подгруппы с наиболее однородными значениями целевой переменной. Затем процесс повторяется для каждой подгруппы до тех пор, пока не будет достигнут критерий остановки (например, достижение определенной глубины дерева, минимального числа элементов в листе и т.д.).

Верхний узел дерева — это корневой узел, определяющий условие тестирования, результат которого соответствует ветви, ведущей к внутреннему узлу. Конечные узлы дерева присваивают классификации (в нашем случае метки мошенничества) и также называются выходными узлами. Каждое дерево также может быть представлено в виде набора правил, поскольку каждый путь от корневой заметки к конечному узлу составляет простое правило «если-то».

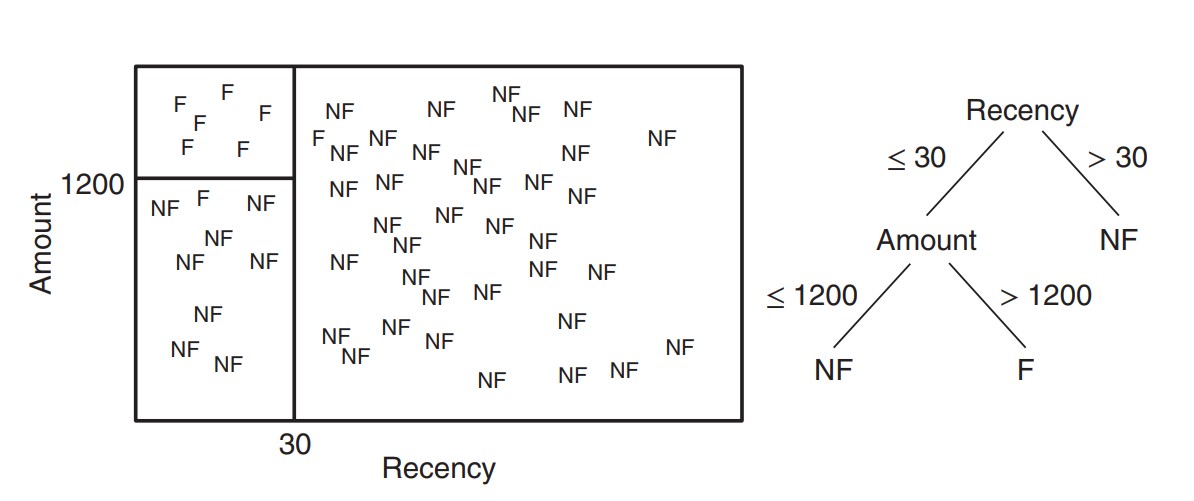


Рисунок 2.7. Границы принятия решений дерева решений.

Деревья решений широко используются в различных областях, таких как бизнес, медицина, финансы, технологии и т.д., для решения задач классификации и регрессии. Они относительно легко интерпретируются и позволяют выявлять важные признаки, которые влияют на целевую переменную. Однако они могут быть склонны к переобучению, особенно при использовании большого числа признаков, и могут быть неэффективными в задачах с большим количеством классов или регрессионных значений.

**Анализ и оценка эффективности существующих решений на основе искусственного интеллекта для борьбы с мошенничеством в банковской сфере**

**Системы обнаружения мошенничества Сбербанка.** В ранний период кибермошенничества в России, Сбербанк использовал набор правил для обнаружения мошеннических транзакций. Эти правила были разработаны экспертами банка, которые регулярно анализировали случаи мошенничества. Каждое правило состояло из нескольких жестких условий, при выполнении которых транзакция считалась подозрительной.

В настоящее же время применяют методы обучения ансамбля классификаторов, которые успешно проявили себя во многих задачах. Прежде чем обучать какую-либо модель, необходимо подготовить данные, на которых модель будет обучаться. У банковских транзакций есть особенность – количество мошеннических транзакций сильно меньше легитимных, что негативно сказывается на качестве модели. В связи с этим, применяются следующие подходы:

1. Передискретизация данных, которая может включать в себя удаление части примеров из основного класса (undersampling), добавление новых примеров в минорный класс путем их дублирования или создания искусственных примеров (oversampling using SMOTE).
2. Методы, основанные на расстоянии, используют прореживание основного класса, однако это делается с учетом расстояний до границ классов, а также путем удаления шумовых или граничных примеров каждого из классов. Примеры таких алгоритмов — Tomek link, One Sided Selection, Neighborhood Cleaning Rule.

Один из применяемых специалистами Сбербанка методов - BalanceCascade - состоит в последовательном обучении серии классификаторов (random forest). На каждом этапе с помощью метода частичной выборки (undersampling) формируется более сбалансированная выборка и обучается классификатор, после чего из исходного набора данных удаляются корректно классифицированные сэмплы основного класса, и операция повторяется вновь. В конечном итоге модель представляет собой набор таких классификаторов.

Как утверждают эксперты Сбербанка[8], переход от статических правил к моделям, построенным с помощью машинного обучения, оказал колоссальное положительное влияние:

* Уровень мошенничества удалось сократить в несколько раз при постоянном росте транзакционной активности и появлении новых продуктов и услуг;
* Автоматизирован процесс реагирования на модификации и появления новых схем мошенничества, увеличена скорость реакции;
* Улучшен клиентский опыт за счет снижения ложных срабатываний системы, а также повышены лимиты проведения операций;
* Процессы оценки эффективности работы стали прозрачнее.

**Система Gotcha!** Здесь рассматривается практическое применение сетевого анализа для обнаружения мошенничества на примере государственных налоговых данных, где компании стремятся уклониться от уплаты налогов. Компании используют различные ресурсы, такие как покупатели, поставщики, заказчики, сотрудники, техника, бухгалтеры и т.д. для своей деятельности. Пример социальной сети предоставлен на рисунке 2.8.

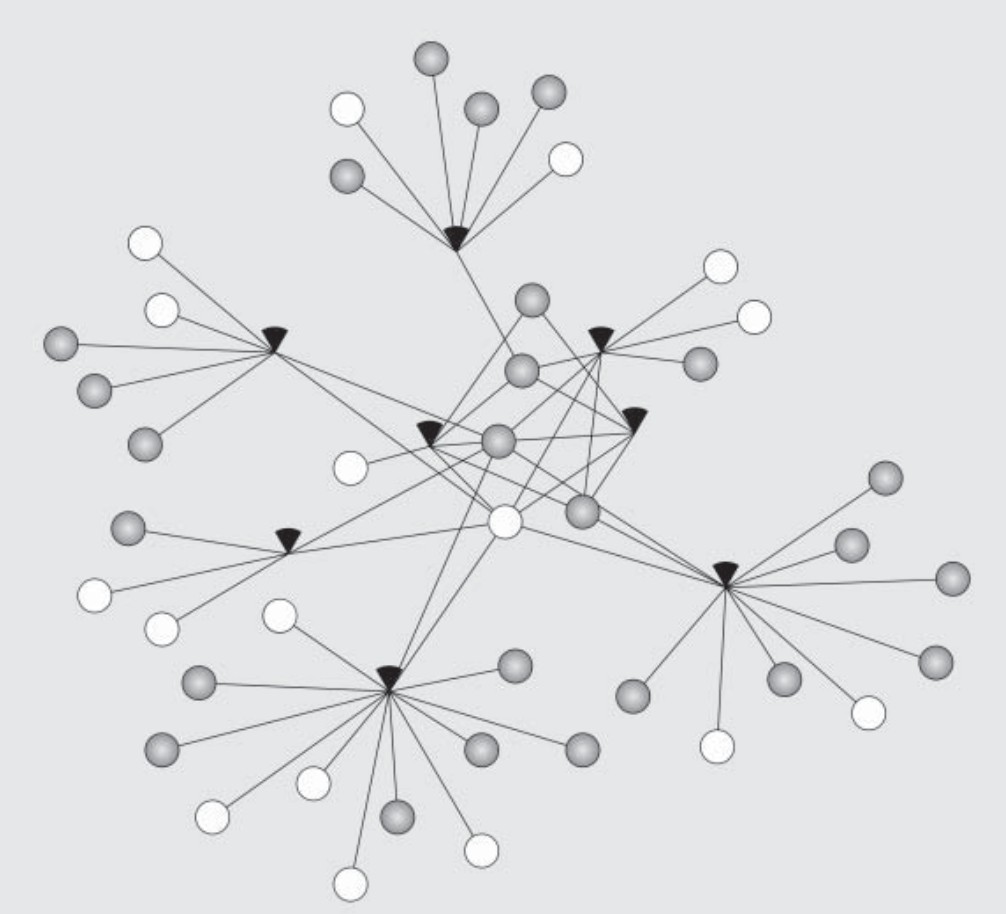


Рисунок 2.8. Образец сети Gotcha!

Компании и ресурсы представлены на графике в виде кружков и клиньев соответственно, причем это двудольная сеть, т.е. объединяет два типа узлов. Вместо того, чтобы напрямую связывать компании друг с другом через общие ресурсы, ресурсы включены в сеть как отдельные объекты. Такой подход позволяет более точно отразить реальность и предоставляет больше деталей об интенсивности связей между компаниями и их ресурсами. Также замечено, что некоторые ресурсы играют важную роль в возникновении мошенничества.

По мнению экспертов, компании строят незаконные конструкции, чтобы избежать уплаты налогов, и тесно сотрудничают для совершения мошенничества. Как правило, компании, являющиеся частью мошеннической схемы, организованы таким образом, что они не получают достаточно прибыли, чтобы погасить свою налоговую задолженность перед государством. Реальную прибыль урезают другие компании, участвующие в установке. Как только компания объявляет о банкротстве и не может продолжать свою деятельность, ресурсы компании перемещаются в пользу других существующих компаний или вновь созданных компаний. Использование сетевого анализа позволяет проследить путь ресурсов от одной обанкротившейся компании к другой, чтобы выявить мошеннические установки в сети. Эти мошеннические установки являются так называемыми «конструкциями пауков». Компании, занимающиеся созданием «пауков», образуют мошенническую сеть и тесно связаны друг с другом посредством ресурсов, которыми они делятся или передают. Тогда перед нами стоит следующая задача: учитывая сеть компаний и ресурсов, как мы можем использовать ярлык нескольких подтвержденных мошеннических компаний, чтобы сделать вывод о вероятности мошенничества для всех других (но законных) компаний в сети.

Gotcha! - это метод выявления мошенничества, который объединяет данные из разных источников, включая изолированную среду компании и отношения между другими компаниями. Этот метод сочетает в себе три типа характеристик данных:

1. Внутренние (или локальные) характеристики: эти характеристики включают в себя характеристики каждого наблюдения (здесь: компания), как если бы наблюдение рассматривалось изолированно. Примеры характеристик включают возраст компании, сектор, в котором работает компания, финансовую отчетность, адрес и т. д.
2. Прямые сетевые характеристики: эти характеристики характеризуют непосредственное соседство или эгонет компании. Напомним, что сеть социального обеспечения представляет собой двудольный граф компаний и ресурсов. Непосредственным соседом компании является компания вместе с ее текущими (и ее прошлыми) ресурсами. Обратите внимание, однако, что мошенничество приписывается только компаниям и что ресурсы и (пока) законные компании изначально не помечены. Косвенные сетевые функции (см. ниже) извлекают первичный признак мошенничества каждого сетевого объекта.
3. Косвенные сетевые характеристики: алгоритм распространения используется для определения вероятности мошенничества для каждого немаркированного ресурса и компании. Алгоритм распространения основан на Google PageRank (см. раздел Алгоритм PageRank) и рассматривает мошенничество как вирус, перемещающийся по сети. Узел в сети, который сильно подвержен мошенничеству, получает высокий балл.

На рисунке 2.9 предоставлена оценка эффективности данного метода.

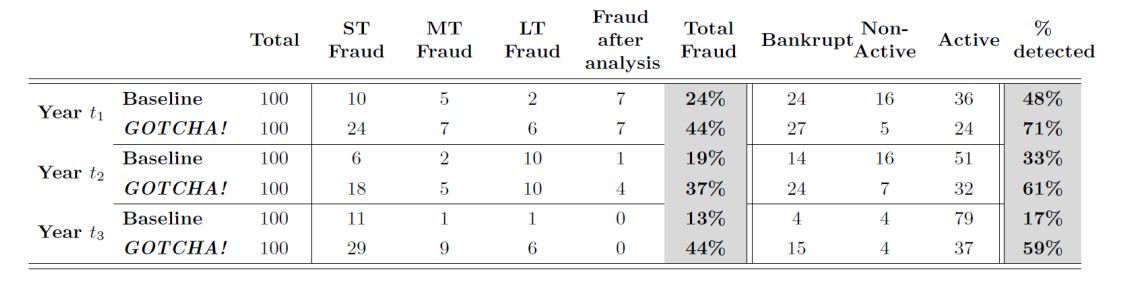


Рисунок 2.9. Эффективность Gotcha! в области социального обеспечения.

**3. ПРОЕКТНЫЙ РАЗДЕЛ**

**3.1 Описание предметной области**

Предметной областью является система обнаружения мошенничества, основанная на алгоритмах машинного обучения.

Система обнаружения мошенничества в банковской сфере - это комплекс программных и аппаратных средств, который осуществляет непрерывный мониторинг и анализ транзакций, связанных с использованием банковских карт, с целью выявления мошеннических операций.

Предметная область системы включает в себя следующие компоненты:

1. Банковские карты: система основана на анализе операций, связанных с использованием банковских карт. Это может включать в себя анализ транзакций, снятия наличных, онлайн-покупок и других операций, связанных с использованием банковских карт.
2. Пользователи банковской системы: пользователи проводят в системе различные операции, например перевод денежных средств на счёт другого пользователя и т.д.
3. Механизм обнаружения мошеннических транзакций. Он включает в себя анализ данных, осуществляя непрерывный мониторинг и анализ данных, связанных с использованием банковских карт (это может включать в себя анализ сумм транзакций, места, даты и времени операций, а также другие факторы), и модель искусственного интеллекта для выявления неугодных операций (машинное обучение, нейронные сети, алгоритмы глубокого обучения и другие методы).

Цель системы обнаружения мошенничества в банковской сфере - обеспечение безопасности финансовых операций клиентов банка и предотвращение мошеннических операций, которые могут причинить финансовый ущерб.

[**3.2. Разработка модели искусственного интеллекта для противодействия мошенничеству в банковской сфере**](#_Toc102888874)

Для разработки алгоритмов искусственного интеллекта я выбрал язык Python, потому что, во-первых, из-за его простоты и удобства, и, во-вторых, из-за богатого набора инструментов для работы с данными, визуализации, предобработки данных и множества библиотек и фреймворков для машинного обучения и глубокого обучения, таких как TensorFlow, PyTorch, scikit-learn, Keras. Все используемые библиотеки перечислены на рисунке 3.1.

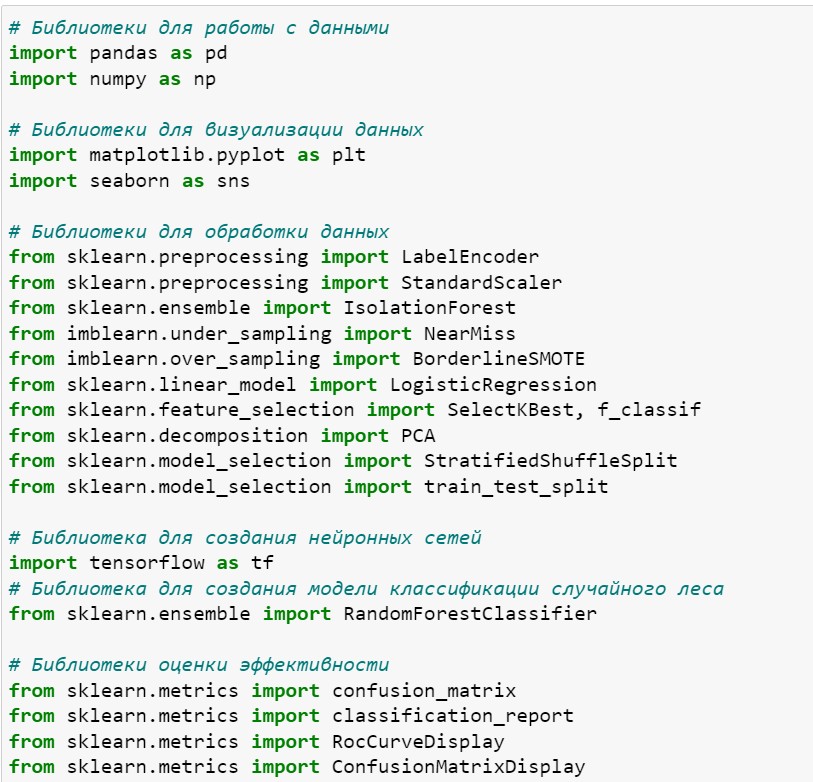


Рисунок 3.1. Используемые Python библиотеки.

Работа с данными будет происходить с помощью библиотеки Pandas, в которой все данные будут хранится в структуре данных, именуемой DataFrame. Данная структура предоставляет удобный интерфейс для наиболее часто необходимых операций, например, объединения двух разных структур или подсчёт количества уникальных значений.

Сами данные операций с кредитными картами взяты из открытого источника[9]. Признакам в этом датасете выданы зашифрованные имена в связи с защитой персональных данных пользователей.

**Список литературных источников**

1. Fraud: Detection, Prevention, and Analytics! [Электронный ресурс] URL: https://www.researchgate.net/publication/315755168\_Fraud\_Detection\_Prevention\_and\_Analytics.
2. Insurance Fraud: Overview, Types of Schemes. [Электронный ресурс] URL: <https://www.investopedia.com/terms/i/insurance-fraud.asp>.
3. What is corruption? [Электронный ресурс] URL: <http://www.corruptie.org/en/corruption/what-is-corruption/>
4. Обзор операций, совершенных без согласия клиентов финансовых организаций – Банк России [Электронный ресурс] URL: <https://www.cbr.ru/analytics/ib/operations_survey_2022/>
5. Третьяк М. И. Различные подходы к оценке способов хищений безналичных денежных средств в условиях современного информационного общества // Всероссийский криминологический журнал. – 2020. – Т. 14, № 4. – С. 601–612. – DOI: 10.17150/2500-4255.2020.14(4).601-612.
6. Криминологический анализ современного состояния мошенничеств в банковской сфере России - А. Л. Репецкая, Л. А. Петрякова [Электронный ресурс] URL: <https://cyberleninka.ru/article/n/kriminologicheskiy-analiz-sovremennogo-sostoyaniya-moshennichestv-v-bankovskoy-sfere-rossii/viewer>
7. [Baesens et al] Fraud analytics using descriptive, predictive, and social network techniques - a guide to data science for fraud detection (2015).
8. Машинное обучение против мошенничества в банковской сфере [Электронный ресурс] URL: <https://www.it-world.ru/cionews/security/118786.html>
9. Credit Card Fraud Detection [Электронный ресурс] URL: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud?datasetId=310&searchQuery=n

**Приложение 1**

*# Библиотеки для работы с данными*import pandas as pd  
import numpy as np  
  
*# Библиотеки для визуализации данных*import matplotlib.pyplot as plt  
import seaborn as sns  
  
*# Библиотеки для обработки данных*from sklearn.preprocessing import LabelEncoder  
from sklearn.preprocessing import StandardScaler  
from sklearn.ensemble import IsolationForest  
from imblearn.under\_sampling import NearMiss  
from imblearn.over\_sampling import BorderlineSMOTE  
from sklearn.linear\_model import LogisticRegression  
from sklearn.feature\_selection import SelectKBest, f\_classif  
from sklearn.decomposition import PCA  
from sklearn.model\_selection import StratifiedShuffleSplit  
from sklearn.model\_selection import train\_test\_split  
  
*# Библиотека для создания нейронных сетей*import tensorflow as tf  
*# Библиотека для создания модели классификации*from sklearn.ensemble import RandomForestClassifier  
  
*# Библиотеки оценки эффективности*from sklearn.metrics import confusion\_matrix  
from sklearn.metrics import classification\_report  
from sklearn.metrics import RocCurveDisplay  
from sklearn.metrics import ConfusionMatrixDisplay

*# Загружаем данные из файла*df = pd.read\_csv("datasets/creditcard.csv")

*# Вывод первых пяти значений*df.head(5)

print("Имена признаков: ", df.columns)  
print("Размерность данных: ",df.shape)

*# Выясним, есть ли пропущенные значения*df.isnull().any().any()

*# Выясним, сколько случаев мошенничества зарегистрировано*df["Class"].value\_counts()

df.describe()

*# Тип данных признаков*df.dtypes

*# Преобразуем категориальные признаки к численным*for column in df.columns:  
 if df[column].dtype == "object":  
 df[column] = LabelEncoder().fit\_transform(df[column])  
 if df[column].dtype == "bool":  
 df[column] = df[column].apply(lambda x: int(x))

*# Отрисовываем диаграмму размаха для всех признаков*plt.figure(figsize=(25, 80))  
  
for i,col in enumerate(df.columns.to\_list()):  
 plt.subplot(int(np.ceil(df.columns.size)), 4, i + 1)  
 sns.boxplot(y=col, data=df)  
  
plt.tight\_layout()  
  
plt.show()

*# Разделим данные на зависимую и независимые переменные*features = df.drop(columns=["Class"]).to\_numpy()  
labels = df["Class"].to\_numpy()

*# Выявим выбросы в данных с помощью алгоритма изоляционного леса*isoForest = IsolationForest(contamination=0.07,random\_state=0)

isoForest.fit(features)

*# выявим выбросы в данных*is\_anomaly = pd.DataFrame(isoForest.predict(features))  
is\_anomaly.value\_counts()

*# Запишем индексы выбросов в ввиде массива булевых значений*anomaly\_indices = list()  
  
for i in range(is\_anomaly.size):  
 *# Оставим данные минорного класса, так как, скорее всего, такие данные были ошибочно  
 # помечены как выбросы* anomaly\_indices.append(is\_anomaly[0][i] == -1 and df["Class"][i] == 0)  
  
anomaly\_indices = df[np.array(anomaly\_indices)].index  
anomaly\_indices[-5:] *# пять последних элементов*

anomaly\_indices.size

*# Доля аномалий относительно всего датасета*anomaly\_indices.size / df.shape[0]

*# Удалим выбросы из датасета, сбросим индексы*df.drop(index=anomaly\_indices, inplace=True)  
df.reset\_index()  
df.head(3)

*# Вновь разделим данные на зависимую и независимые переменные*features = df.drop(columns=["Class"]).to\_numpy()  
labels = df["Class"].to\_numpy()

*# Выберем 10 лучших признаков с помощью алгоритма selectKBest   
# на основании значения ANOVA-F*kBest = SelectKBest(score\_func=f\_classif, k=10)

fitResult = kBest.fit(features, labels)  
print("Оценки признаков: ", fitResult.scores\_)  
print("Индексы лучших признаков:", kBest.get\_support(indices=True))

*# Достанем имена признаков из номеров*most\_important\_features\_names = df.columns[kBest.get\_support(indices=True)].to\_list()  
print("Имена 10 самых значимых признаков", most\_important\_features\_names)  
*# Выберем 10 самых значимых признаков*selected\_features = features[:, kBest.get\_support(indices=True)]  
selected\_features[:3]

*# Стандартизируем данные*std\_scaler = StandardScaler()  
selected\_features = std\_scaler.fit\_transform(selected\_features)  
selected\_features[:2]

selected\_features.shape, labels.shape

*# Разделим данные на обучающую и тестовую выборки в отношении 7/3  
# особенность разделения в том, что данные делятся на основании   
# значений зависимой переменной, т.е. в обучающей выборке гарантированно   
# будет 70% мошеннических случаев и в тестовой 30%*sss = StratifiedShuffleSplit(n\_splits=8, test\_size=0.3, random\_state=0)  
sss.get\_n\_splits(selected\_features, labels)

df\_features = df.drop(columns=["Class"]).copy()

*# Запомним индексы выбранных значений*train\_indices, test\_indices = list(), list()  
  
for i, (train\_index, test\_index) in enumerate(sss.split(df\_features, labels)):  
 print(f"№ сортировки: {i}:")  
 print(f" Train: index={train\_index}, shape={train\_index.shape}")  
 print(f" Test: index={test\_index}, shape={test\_index.shape}")  
 train\_indices, test\_indices = train\_index, test\_index

selected\_features.shape

X\_train, X\_test = selected\_features[train\_indices], selected\_features[test\_indices]  
Y\_train, Y\_test = labels[train\_indices], labels[test\_indices]

*# Доля мошеннических случаев в тестовой и тренировочной выборках*print("Доля мошеннических случаев в тренировочной выборке = ", np.sum(Y\_train == 1))  
print("Доля мошеннических случаев в тестовой выборке = ", np.sum(Y\_test == 1))

*# Выясним долю мошеннических операций*print(f"Доля мошеннических операций = {(np.sum(Y\_train == 1) / np.sum(Y\_train == 0)) \* 100} %")

*# Выкинем данные по алгоритму NearMiss для увеличения доли мошеннических операций до 5%*underX\_Train, underY\_Train = NearMiss(sampling\_strategy=0.03).fit\_resample(X\_train, Y\_train)

underX\_Train.shape

*# Выясним долю мошеннических операций*print(f"Доля мошеннических операций = {(np.sum(underY\_Train == 1) / np.sum(underY\_Train == 0)) \* 100} %")

print(np.sum(underY\_Train == 0))  
print(np.sum(underY\_Train == 1))

*# Сгенерируем данные мошеннических операций алгоритмом SMOTE, подвидом borderline-1*smoteAlg = BorderlineSMOTE(kind='borderline-1', sampling\_strategy="minority")  
preparedX\_train, preparedY\_train = smoteAlg.fit\_resample(underX\_Train, underY\_Train)

preparedX\_train.shape

print(np.sum(preparedY\_train == 0))  
print(np.sum(preparedY\_train == 1))

**Приложение 2**

feature\_normalization\_layer = tf.keras.layers.Normalization()  
feature\_normalization\_layer.adapt(preparedX\_train)

*# Объявим модель с входным слоем, 5-ю скрытыми слоями с фун-ей активации selu  
# и выходной слой с фун-ей активации sigmoid для бинарной классификации*model = tf.keras.Sequential([  
 *#feature\_normalization\_layer,* tf.keras.layers.Input(shape=(preparedX\_train.shape[1],)),  
 tf.keras.layers.Dense(units=128, activation="selu", kernel\_initializer="lecun\_normal"),  
 tf.keras.layers.Dense(units=256, activation="selu", kernel\_initializer="lecun\_normal"),  
 tf.keras.layers.Dense(units=512, activation="selu", kernel\_initializer="lecun\_normal"),  
 tf.keras.layers.Dense(units=256, activation="selu", kernel\_initializer="lecun\_normal"),  
 tf.keras.layers.Dense(units=128, activation="selu", kernel\_initializer="lecun\_normal"),  
 tf.keras.layers.Dense(units=1, activation="sigmoid")  
])  
  
model.summary()

*# Соберем модель с функцией потерь бинарной кроссэнтропией,   
# Оптимизатором Адам и метриками полноты и бинарной точности для оценки эффективности*model.compile(  
 loss=tf.keras.losses.BinaryCrossentropy(),  
 optimizer=tf.keras.optimizers.Adam(learning\_rate=10\*\*-4),  
 *#optimizer=tf.keras.optimizers.SGD(),* metrics=[tf.keras.metrics.Recall(), tf.keras.metrics.BinaryAccuracy()]  
)

*# Начнем обучать модель на преобразованных данных 100 эпох, делая проверки на тестовой выборке каждые две эпохи  
# Установим фунцию ранней остановки обучения, которая прервет обучение, если полученная   
# метрика "val\_binary\_accuracy" будет 4 раза снижаться, после остановки модели будут присвоены веса на   
# лучшей эпохе*model.fit(preparedX\_train, preparedY\_train, epochs=100, batch\_size=32 ,  
 callbacks=tf.keras.callbacks.EarlyStopping(monitor="val\_binary\_accuracy", patience=4, restore\_best\_weights=True),  
 validation\_data=(X\_test, Y\_test), validation\_freq=2)

y\_predict = model.predict(X\_test)

*# Получим предсказания нейронной сети на тестовой выборке*nn\_y\_predict = model.predict(X\_test)

*# Преобразуем вероятности в классы*y\_predict\_labels = np.array(np.floor(y\_predict + 0.5), dtype=np.int64)  
y\_predict\_labels

*# Построим классификатор случайного леса и обучим его*randomForest = RandomForestClassifier(random\_state=5, n\_estimators=100, n\_jobs=-1)  
randomForest.fit(preparedX\_train, preparedY\_train)

*# Получим предсказания случайного леса*rndFrst\_predicted = randomForest.predict(X\_test)

**Приложение 3**

*# Отрисуем кривые обучения*plt.figure(figsize=(12, 7))  
  
plt.plot(history.epoch, history.history['loss'], label="Потери на тренировочной выборке")  
plt.plot(history.epoch[1::2], history.history['val\_loss'], label="Потери на тестовой выборке")  
  
plt.title('Кривые обучения для нейронной сети в бинарной классификации мошенничества', size=20)  
plt.xlabel('Номер эпохи', size=14)  
  
plt.legend();

*# Получим отчет о классификации нейронной сети*print('Classification report:\n', classification\_report(Y\_test, nn\_y\_predict\_labels))

*# Построим матрицу ошибок*nn\_conf\_mat = confusion\_matrix(Y\_test, nn\_y\_predict\_labels)  
nn\_conf\_mat

*# Отобразим матрицу ошибок нейронной сети*names = ["True Positives", "False Positives", "False Negatives", "True Negatives"]  
label\_names = list()  
  
for i in range(len(names)):  
 label\_names.append(names[i] + "\n" + str(nn\_conf\_mat.reshape(-1)[i]))  
  
label\_names = np.asarray(label\_names).reshape(2,2)  
sns.heatmap(nn\_conf\_mat, annot=label\_names, fmt="", cmap="Blues")

*# Построим ROC кривую*RocCurveDisplay.from\_predictions(y\_true=Y\_test, y\_pred=nn\_y\_predict\_labels)

*# Получим отчет о классификации*print('Classification report:\n', classification\_report(Y\_test, rndFrst\_predicted))

*# Вычислим матрицу ошибок случайного леса*rnd\_frst\_conf\_mat = confusion\_matrix(Y\_test, rndFrst\_predicted)  
rnd\_frst\_conf\_mat

*# Отобразим матрицу ошибок случайного леса*names = ["True Positives", "False Positives", "False Negatives", "True Negatives"]  
label\_names = list()  
  
for i in range(len(names)):  
 label\_names.append(names[i] + "\n" + str(rnd\_frst\_conf\_mat.reshape(-1)[i]))  
  
label\_names = np.asarray(label\_names).reshape(2,2)  
sns.heatmap(rnd\_frst\_conf\_mat, annot=label\_names, fmt="", cmap="Blues")

*# Построим ROC кривую для случайного леса*RocCurveDisplay.from\_predictions(y\_true=Y\_test, y\_pred=predicted)