**Распознавание мошеннических операций в мобильных банковских системах с использованием алгоритмов машинного обучения, и информирование пользователей о рисках**

***Скуратов А.Ю.***

*Студент*

*Сибирский Государственный Университет Путей Сообщения, Факультет Бизнес-Информатики, Новосибирск, Россия*

*E-mail: skuratov.1411@gmail.com*

Актуальность проблемы мошенничества можно определить, взглянув на статистику. По данным Центробанка за 2019 год мошенники украли у банковских клиентов 6 млрд рублей. При этом, кредитные организации вернули всего 14% от данной суммы. Идём дальше, в 2020 году банки вернули уже 11%, из похищенных 9 млрд. А в 2021 6% из 13 млрд. Закономерность очевидна: падение уровня возврата средств и стремительный рост объема хищений.

Машинное обучение – это стремительно развивающаяся область информатики. Оно находит множество сфер применения, например, распознавание заболеваний в медицине, классификация квантовых частиц в физике, фильтрация почты, системы рекомендаций, компьютерное зрение и так далее. Машинное обучение получило огромный скачок популярности в последнее время. С ростом количества информации и данных, машинное обучение будет становиться всё более мощным инструментом.

Цель данного проекта: разработать приложение, имитирующие мобильный банк, и внедрить туда систему распознавания мошеннических транзакций и уведомления о них пользователей, работающую с помощью методов машинного обучения.

Для достижения поставленной цели, потребовалось решить следующие задачи:

1. исследовать данные о транзакциях;
2. выбрать, настроить и обучить алгоритм машинного обучения;
3. разработать приложение имитирующие мобильный банк;
4. добавить визуализацию риска мошеннической транзакции;
5. реализовать уведомление пользователя.

Как было сказано ранее, машинному обучению нужны данные, на которых мы будем обучать алгоритм. Для настоящего проекта в свободном доступе был найден набор данных (датасет), содержащий информацию о 300 тысяч транзакций.

Перед поиском алгоритма, следует провести некоторый анализ данных, для возможного преобразования датасета. Благодаря этому получится оставить самые информативные данные, что, несомненно, пойдёт на пользу алгоритму. В данной работе, из-за особенностей датасета, наиболее эффективным оказался метод under-sampling, или метод под выборки, который предназначен для борьбы с дисбалансом классов.

После анализа производится поиск алгоритма машинного обучения. Обычно выбирается с десяток различных алгоритмов, каждый из которых тщательно настраивается и тестируется. Таким образом ведется поиск самого точного алгоритма. В настоящей работе были протестированы такие алгоритмы, как: наивный байесовский классификатор, логистическая регрессия, градиентный бустинг и др. Сравнив точность каждого алгоритма, лучшим оказался алгоритм под названием случайный лес (Random Forest). Это так называемый ансамбль алгоритмов, состоящий из большого количества решающих деревьев. Каждое дерево классифицирует транзакцию как мошенническую или не мошенническую, а итоговым ответом алгоритма, будет соотношение ответов деревьев, или иначе, вероятность транзакции принадлежать к классу мошеннических. Алгоритм способен распознать мошенническую транзакцию, с вероятностью в 93%.

Приложение на телефон было разработано с помощью языка программирования Kotlin. Для хранения данных о транзакциях было использовано облачное хранилище данных FireBase. В приложении были реализованы: процесс совершения транзакций, отображение истории операций, уведомление пользователей о рисках мошенничества, графическая визуализация рисков транзакций.

При совершении транзакции, данные по ней отправляются на сервер и обрабатываются алгоритмом. После этого приложение считывает вероятность мошеннической транзакции и сохраняет в истории и на графике. Если вероятность превышает 70% пользователю придёт уведомление о риске мошеннической транзакции.

По итогу, была разработана система распознавания мошеннических транзакций, которая работает с высокой точностью в 93%. Уведомления о мошеннических транзакциях, поможет более оперативному обращению пользователей в банк, как и отображение рисков транзакций в истории и на графике. Возможны и изменения данной системы: за место того, чтобы просто уведомлять пользователя о рисках, алгоритм может отклонять подозрительные транзакции или, к примеру, спрашивать подтверждения пользователя.