Использование моделей машинного обучения для выявления подозрительных операций в деятельности кредитных организаций

Выполнила: Беспалова Ольга Александровна

Руководитель: к.т.н. Семендяев Родион Юрьевич

Термины и определения

Подозрительные операции - операции с денежными средствами или иным имуществом, предположительно совершаемые в целях легализации (отмывания) доходов, полученных преступным путем, и финансирования терроризма - 115-ФЗ от 7 августа 2001 года «О противодействии легализации (отмыванию) доходов, полученных преступным путём, и финансированию терроризма»

Затраты на соблюдение требований законодательства в сфере ПОД/ФТ

- Размер штрафов за несоблюдение ПОД/ФТ в мире составил порядка $1 трлн (в 2021г.).
- Расходы на соблюдение ПОД/ФТ в мире оцениваются примерно в $10 трлн в год.
- Расходы на соблюдение законодательства в сфере под/фт - одна из основных статей расходов как для российских, так и для зарубежных банков.

Цель работы

Оценить возможность выявления подозрительных операций методами машинного обучения.

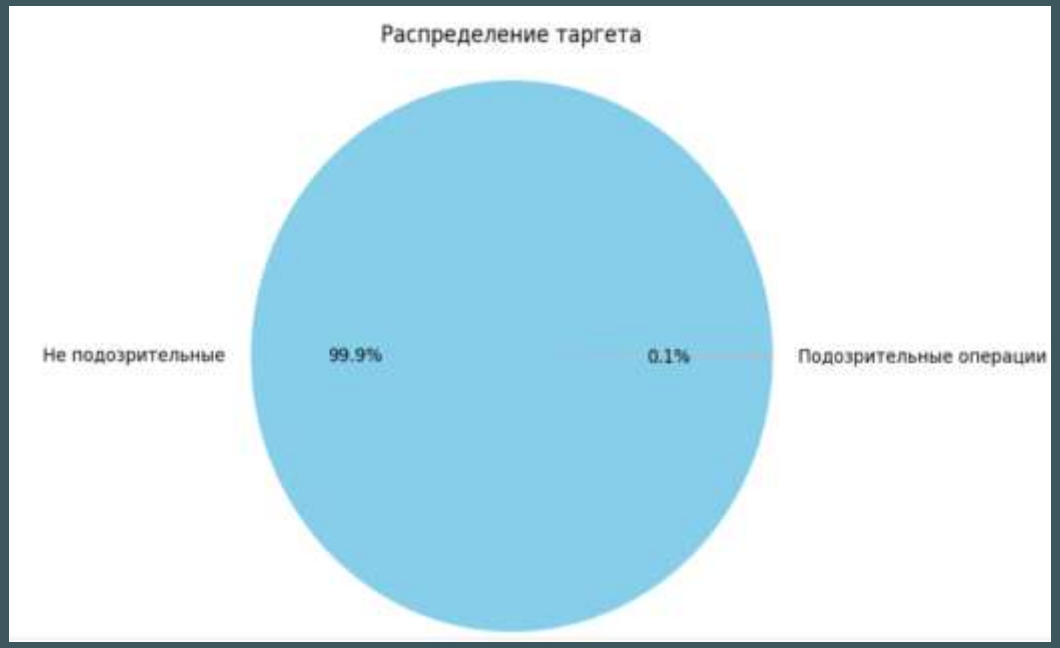
Задачи работы

1. Поиск данных для проведения исследования.
2. Проведение первичного анализа и подготовка данных.
3. Построение моделей машинного обучения.
4. Анализ метрик, выбор лучшей модели.

Данные (датасеты)

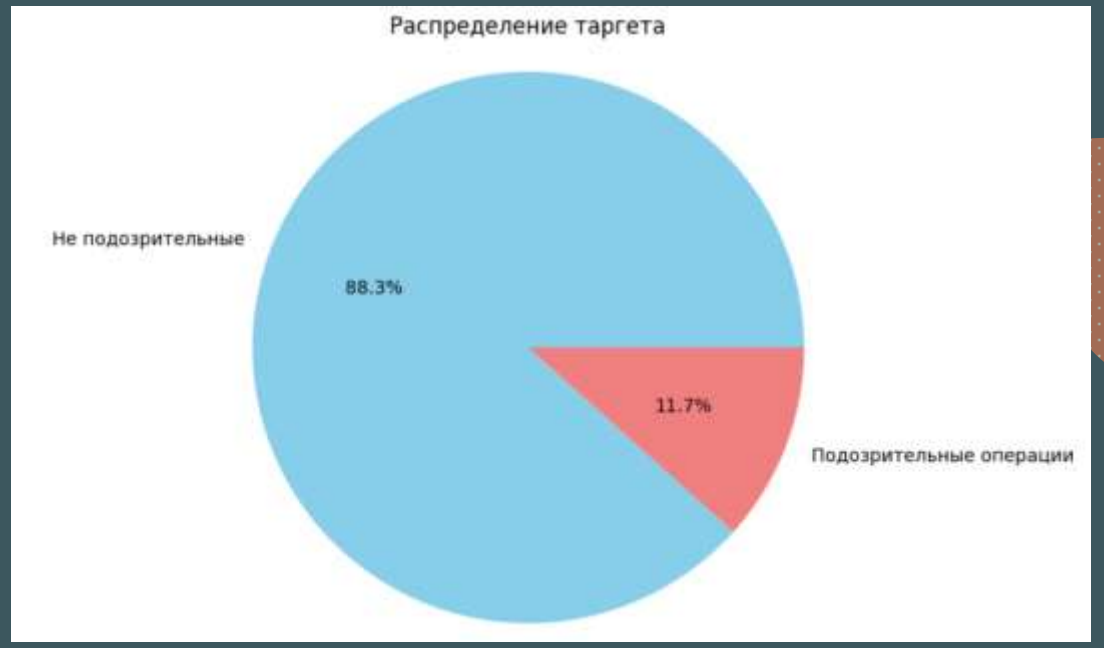
Датасет №1

- 6924049 наблюдения;
- 11 признаков
- Целевая переменная (таргет) – подозрительность транзакции



Датасет №2

- 1424 наблюдения;
- 33 признака
- Целевая переменная (таргет) – подозрительность транзакции



Работа с признаками

- Проверка на пропущенные значения и дубликаты
- Генерирование новых признаков из существующих
- Оценка значимости признаков
- Удаление избыточных признаков
- Преобразование значений признаков в формат, пригодный для машинного обучения (кодирование категориальных признаков, нормализация).

Количество признаков после преобразования

Датасет 1

11 → 12

Датасет 2

33 → 21

Применяемые подходы для построения моделей машинного обучения в рамках ВКР

- Обучение с учителем.
- Обучение без учителя.

Модели обучения с учителем

- Логистическая регрессия;
- Наивный Байесовский классификатор;
- Дерево решений;
- Случайный лес;
- Метод К-ближайших соседей;
- Градиентный бустинг;

Модели обучения без учителя

- Изолирующий лес;
- Автокодировщики;
- Одноклассовый метод опорных веторов.

Метрика оценки качества моделей

F1 - гармоническое среднее значение точности (precision) и полноты (recall).

Значение F1 – от 0 до 1.

Точность (precision) - сколько реальных объектов класса среди всех тех, что классификатор отнёс к этому классу.

Полнота (recall) - какую долю из общих объектов класса составляют найденные объекты.

Метрики моделей машинного обучения с учителем для датасета 1

	LR	DC	RF	KNN	NB	GS
F1	**0.05**	**0.51**	**0.61**	**0.36**	**0.13**	**0.59**

Метрики моделей машинного обучения без учителя для датасета 1

	Autoencoder	One-Class SVM	Isolation Forest
F1	**0.58**	**0.14**	**0.1**

Понижающий отбор (undersampling)

- **Undersampling** — это метод уменьшения выборки, который используется для достижения баланса между классами в несбалансированных данных.
- **Принцип –** случайное удаление части примеров из большего класса до достижения желаемого баланса.

Метрики моделей машинного обучения с учителем для датасета №1 с применением понижающего отбора.

	LR	DC	RF	KNN	NB	GS
F1	**0.33**	**0.51**	**0.48**	**0.37**	**0.1**	**0.45**

Метрики моделей машинного обучения с учителем для датасета 2

	LR	DC	RF	KNN	NB	GS
F1	**0.9**	**0.97**	**0.97**	**0.88**	**0.88**	**0.98**

Метрики моделей машинного обучения без учителя для датасета 2

	Autoencoder	One-Class SVM	Isolation Forest
F1	**0.75**	**0.12**	**0.53**

Выводы

- Модели, обученные на размеченных данных (обучение с учителем), показали метрики выше , чем модели, обученные на неразмеченных данных (обучение без учителя).
- На данных датасета № 2 модели машинного обучения показали метрики значительно выше, чем на данных датасета №1.
- Разница в метриках может быть объяснена лучшим качеством имеющихся исходных данных датасета № 2, наличием более информативных признаков, которые помогают моделям машинного обучения более эффективно выявлять закономерности и обучаться.

Выводы

- Выявление подозрительных операций кредитными организациями методами машинного обучения потенциально может применяться успешно, однако, успешность такого подхода к обнаружению подозрительных транзакций зависит от качества и полноты данных, на которых обучаются модели машинного обучения, а также от выбора лучшей модели для каждого набора данных.