**Эссе по кейсу.**

Астраханцев Роман, СКБ-171

Говорить хороша модель или плоха – это не совсем корректно. Всё зависит от задачи, решаемой, данной модели и действий, которые мы будем совершать по результатам работы модели. А именно, если мы не желаем лишний раз беспокоить пользователей, совершающих легитимные операции, то важна точность определения мошеннических транзакций. В таком случае предоставленная модель будет иметь много лишних срабатываний. Если же Сбербанку важнее предотвратить как можно больше мошеннических транзакций, то метрикой качества в данном случае будет являться полнота. Как можно видеть, у модели данный показатель составляет 85%, что довольно много. Итого можно сказать, что данная модель хотя и имеет много ложных срабатываний, но она способна определить почти всех мошенников.

Исходя из полученных данных о точности и полноте и в предположении, что мошеннических операций не происходит более 2000 шт. в сутки, оценим суточную нагрузку на сотрудников, осуществляющих обзвон подозрительных операций. Составим матрицу ошибок и введём в ней обозначения.

Тип транзакции

|  |  |  |
| --- | --- | --- |
|  | Мошенническая | Легитимная |
| Мошенническая  Пометка модели | TP | FP |
| Легитимная | FN | TN |

Сотрудники кол-центра должны будут обзвонить всех пользователей, чьи транзакции были признаны моделью мошеннической (то есть подозрительной). Таким образом величина TP+FP является целевой. Предположение об общем числе мошеннических операций за сутки выражается в FN+TР. Полнота равна TP/(FN+TP) и равна соответственно 0.85, а точность получается как TP/(TP+FP) и равна 0.21. Составив систему уравнений, легко получить, что TP+FP = (FN+TP) \* полнота / точность = 8095.238. Таким образом получили, что в день работникам кол-центра придётся совершить не более 8096 звонков.

Как можно видеть исходя из цифр выше, более 6000 легитимных транзакций в день будут приостановлены. Соответственно прежде всего необходимо улучшить точность распознавания мошенников. Учитывая это, можно также предположить, что будет большое количество транзакций без обратной связи. Просто так использовать в разметке такие транзакции нельзя, поскольку это может быть как легитимная транзакция, которую пользователь повторил чуть позже, либо же это может быть мошенническая транзакция, которую не заметил пользователь. Тривиальным решением данной ситуации будет повторная попытка позвонить пользователю, однако такой метод приведёт к дополнительной нагрузке на, итак, загруженный кол-центр. Другим методом может стать полное исключение подобного рода данных из процесса дообучения модели, однако продолжат фигурировать в итоговых метриках по работе антифрод системы как целое.

Посмотрев на ситуацию внимательнее, можно подумать об автоматической разметки такого класса транзакций. В таком случае транзакцию нельзя рассматривать как моментальное событие, а как одно из цепочки других событий. Так, например, если после какой-то подозрительной транзакции счета пользователя, с которого были сняты средства, были приостановлены, а карты заблокированы, то логичным выводом будет пометить зафиксированную транзакцию как мошенническую. Таким образом, можно создать модель с фичами, которые смотрят не только на саму транзакцию, но и на действия, произошедшие после неё.

Кроме того, для улучшения эффективности модели, стоит так же учитывать некий временной интервал и до операции, то есть если были произведены P2P переводов через сторонние сайты, заблокированные двухфакторной аутентификацией, и после неё была совершен перевод через сайт, не требующий такую проверку, то также будет логично предположить, что мошенники пытались через разные сервисы вывести деньги. В таком случае получится заблокировать транзакцию непосредственно в момент совершения.

Таким образом, можно сказать, что предоставленная модель является неплохой стартовой точкой для распознавания мошеннических транзакций. Есть возможность экспериментирования и проверкой гипотез относительно разметки транзакций без обратной связи.