

## Предварительный анализ задачи определения мошеннических финансовых операций

### Домашнее задание № 1

**Цель работы.** В данном домашнем задании Вы потренируетесь формализовывать поставленные бизнесом задачи анализа данных, определять цели и способы их достижения, декомпозировать систему анализа данных и визуализировать задачи с использованием канбан-доски в GitHub Projects.

Уважаемый слушатель!

Вас пригласили на работу в качестве *специалиста по данным* в коммерческую компанию, занимающуюся предоставлением банковских услуг населению. Одной из услуг компании, пользующейся большой популярностью у населения, является *проведение онлайн-платежей* с банковских счетов граждан. К сожалению, в последнее время участились случаи *мошенничества* при проведении таких операций. Руководство компании встревожено такой ситуацией. На ближайшем совещании было принято решение внедрить в IT-инфраструктуру компании модуль на базе машинного обучения, способный в реальном времени оценивать, является ли проводимая транзакция мошеннической. Это, в свою очередь, позволило бы системе онлайн-платежей отклонить подозрительную операцию. Вам предлагают принять участие в проектировании и построении такой системы.

В ходе детального обсуждения проекта с представителями заказчика выяснились следующие моменты:

1. Руководство готово выделить на реализацию данного проекта не более 10 млн. руб. (не считая зарплат специалистам). Через три месяца оно хотело бы увидеть первые результаты, что позволило бы понять, стоит ли дальше развивать проект. В случае положительного решения, проект нужно завершить за полгода.

2. Из открытых источников известно, что у ближайших конкурентов на каждую сотню проводимых их клиентами транзакций фиксируется не более двух мошеннических, приводящих к потере денежных средств. При этом общий ущерб клиентов за месяц не превышает 500 тыс. руб. Разрабатываемая система должна выдавать результаты не хуже, иначе компания станет неконкурентоспособной.

3. В среднем, компания обрабатывает около 50 транзакций в секунду, однако, перед праздниками это число может достигать 400.

4. Если система определит корректную транзакцию как мошенническую, эта транзакция будет отклонена, а пользователь будет недоволен.

Опыт бизнес-аналитиков подсказывает, что если доля таких транзакций превысит 5 %, то начнется отток клиентов из компании.

5. Представители компании не готовы размещать разрабатываемый модуль на собственных вычислительных ресурсах.

6. При проведении транзакции системой вся информация о ней сохраняется в виде **csv**-файлов, каждая строка которых соответствует одной транзакции. Информация о транзакциях некоторого фиксированного периода помещается в отдельный файл.

7. Файлы с данными содержат всю информацию о транзакциях, включая данные клиента, который ее проводил. Данная информация является конфиденциальной, ее утечка недопустима.

**Вам предлагается** на основе представленной информации:

1. Сформулировать цели проектируемой антифрод-системы в соответствии с требованиями заказчика.

2. Аргументированно выбрать метрику машинного обучения, оптимизация которой, на Ваш взгляд, позволит достичь поставленных целей.

3. Проанализировать особенности проекта с использованием, например, **MISSION Canvas**. Это позволит выделить специфику проекта и формализовать требования к нему, подобрать инструментарий и критерии оценки.

4. Попытаться декомпозировать планируемую систему, определить ее основные функциональные части.

5. Определить задачи, решение которых необходимо для достижения поставленных целей с учетом проведенного анализа. Задачи рекомендуется формулировать по принципу **S.M.A.R.T.** На текущий момент, пока не конкретизированы детали антифрод-системы, они могут быть представлены в достаточно общем виде. Но они позволят сформировать некоторый **roadmap** по реализации проекта.

6. Создать GitHub-репозиторий для проекта и разместить в нем сформулированные задачи (~ 5 задач) на **Kanban**-доске в **GitHub Projects** со статусом «Новые». В дальнейшем, при реализации проекта, эти задачи будут конкретизироваться, дробиться, видоизменяться и выполняться. Их статус тоже будет изменяться, что позволит Вам наглядно оценивать степень завершенности проекта.

**Для получения положительной оценки** за работу необходимо выполнить минимум первые три вышеприведенные задания.

***Желаем успехов!***