Общая идея реализуемого MFDP (выбран кейс 1) – отбор проектов для инвестирования которые в перспективе «взлетят» (соберут таргет + перевыполнят его) на основе опубликованной информации о проекте.

**Бизнес-анализ (Business Understanding)**

* *Явный заказчик отсутствует, основные пользователи – инвесторы фонда*
* *Бизнес-цель проекта заключается в оптимизации инвестиционной стратегии через отбор для финансирования только успешных проектов*
* *В настоящее время на рынке уже есть такие решения как CB Insights, Crunchbase, DataFox, Tracxn, однако это международные компании, которые не предоставляют часть услуг россиянам/российским проектам и ограничивают функционал или имеют высокую стоимость/закрыты для внешних пользователей.*

**1. Текущая ситуация (Assessing current solution)**

Оцениваем, хватает ли ресурсов для проекта.

* Проект и хранение данных будут реализовываться на собственных мощностях + Kaggle Notebook

В части рисков проекта и способов их митигации:

* Нарушить сроки проекта – создание план-проекта для детального планирования предстоящих этапов внедрения.
* Качество данных – анализ ряда альтернативных источников, в том числе не только по платформе кикстартер.
* Отсутствие закономерностей – проработка альтернативных источников для дополнительных фичей.
* Недостаток ресурсов для обучения модели – для экономии времени необходим предварительный анализ имеющихся вариантов решений похожих задач для использования в проекте наиболее подходящих.

**2 Решаемые задачи с точки зрения аналитики (Data Mining goals)**

* Поскольку есть гипотеза о несбалансированности нашей выборки, то в качестве основной метрики проекта будет выбрана F-мера, в качестве дополнительных добавляется PR-AUC и потенциальная прибыль (разница между целевым и фактическим сбором средств)
* Минимальный критерий успешности будет 0.5, целевой 0.7.

План проекта:

1. Проведение бизнес-анализа выбранного кейса (до 26.08.2024)
2. Анализ данных (26.08.2024-29.08.2024)
3. Пауза в связи с отсутствием 😊 (30.08.2024-08.09.2024)
4. Подготовка данных (08.09.2024-14.09.2024)
5. Моделирование и оценка получившихся результатов (14.09.2024-29.09.2024)

Создание «сервиса» для модели (29.09.2024-07.10.2024)