**Шаг 1. Business Understanding**

Понять, что конкретно хочет получить от нас Заказчик. Как он видит результат нашей деятельности. Определить критерии успеха проекта. Насколько точной должна быть наша модель, определяющая вероятность покупки товаров, лежащих в корзине.

Проанализировать наши ресурсы. Сможем ли мы при имеющейся команде сотрудников предоставить заказчику результат работы, отвечающий всем требованиям и ограничениям, либо нужно привлекать дополнительные мощности, трудовые ресурсы (аутсорсинг, краудсорсинг)

Провести анализ рисков. Удастся ли нам добиться той точности модели, которую ожидает от нас Заказчик. Достаточно ли данных в источниках, на основании которых мы будем строить нашу модель.

Провести анализ затрат/выгод. Каковы будут затраты на реализацию данного проекта на наших мощностях, либо при привлечении аутсорса, краудсорса. Какова прогнозируемая прибыль.

После анализа вышеперечисленного, если нас все устраивает, то прикинуть примерный план действий.

**Шаг 2. Data Understanding**

Ввести в курс дела команду. Обсудить все ли имеется у команды для выполнения задачи и какие могут быть непредвиденные обстоятельства.

Обсудить план действий и выбрать инструменты. Какие методы помогут достигнуть максимального качества модели при минимальных затратах. Как будет представлен результат. Необходима ли интеграция нашей модели с интернет магазином заказчика либо он ожидает видеть результат иначе.

Если все хорошо, то начинаем работу.

Собираем данные со всех необходимых источников данных. Описываем их качество и количество. Анализируем данные на ошибки, пропуски, недостатки метаданных.

**Шаг 3. Data Preparation**

Определяем основные атрибуты, гипотезы, множества и подмножества. Определяем стак данных, необходимых для качественного расчета конверсии продаж товаров из корзины. После выбора данных разделяем их на основную и тестовую выборки.

Избавляемся от ошибок, убираем дубликаты, по возможности заполняем пропуски, расширяем данные новыми признаками. Проверяем корреляцию данных с итоговым показателем (Например, как сильно взаимосвязаны история по пользователям и истории их товаров с конверсией корзины в покупку). Определяем приоритетность данных для того, чтобы понять, что конкретно нам использовать.

Сохраняем данные для последующего построения модели.

**Шаг 4. Modeling**

Выбираем метод моделирования (или несколько методов), строим модель, проводим расчеты. Определяем точность модели (если их несколько, то выбираем самую точную, либо при одинаковой точности выбираем ту, которая требует меньше энергозатрат). Проводим эксперимент как работает модель для наших условий. Насколько точна вероятность покупки товаров из корзины, хватает ли заданных источников данных, есть ли ошибки в модели. Описываем результаты работы для дальнейшего анализа.

**Шаг 5. Evaluation**

Проанализировать итоги результатов моделирования. Достигла ли модель требуемых результатов. Какие результаты показывает нам модель и не появились ли в процессе какие-либо новые идеи (привлечение дополнительных данных, смена модели, методы снижения затрат).

Нет ли каких-либо трудностей и дополнительных вопросов. Есть ли какие-нибудь недочеты, стоит ли их исправлять, и если стоит, то как именно это сделать.

Сможем ли мы предоставить результат своей работу заказчику, и сможет ли он его грамотно принять и в дальнейшем беспроблемно с ним работать.

Если все хорошо, то спланировать алгоритм внедрения.

**Шаг 6. Deployment**

Если у нас было несколько моделей, то составляем план внедрения для каждой из них. Определяем для какой модели осуществление внедрения пройдет самым эффективным способом.

Предполагаем возможные риски и проблемы во время внедрения и заранее планируем свои действия при возникновении этих проблем. (Возможны проблемы с ПО у заказчика, несовместимость версий, недостаток ресурсов для обслуживания)

Определяем дальнейшее взаимодействие с заказчиком. Будет ли ему требоваться поддержка для работы с моделью с нашей стороны в дальнейшем. Будет ли требоваться актуализация модели и что для этого придется сделать. Как часто будет требоваться актуализация (Если часто, то может быть стоит переделать нашу модель так, чтобы ее было проще адаптировать к изменяющимся условиям). Будет ли работать модель при обновлении версий ПО.

Провести итоговый обзор проекта.