|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Очередь задач  (To do) | Понимание бизнеса  (Business Understanding) | Изучение источников данных, сбор данных  (Data Understanding) | Подготовка данных  (Data Preparation) | Моделирование (Modeling) | Оценка (Evaluation) | Развертывание (Deployment) |
| 5 max | 2 max | 2 max | 2 max | 2 max | 2 max | 2 max |
| Расставить задачи в порядке приоритета. | 1. Определение бизнес цели т.е. формализация ожиданий заказчика.  2. Математическая постановка задачи анализа данных.  2.1 Выбор оптимизируемой метрики и её значения достаточного для удовлетворения ожиданий заказчика.  2.2 Определение признакового описания объекта.  2.3 Определение списка источников данных из которых признаковое описание может быть получено.  2.4 Определение сроков  Минутки (протоколы, письма) фиксирующие результаты встреч с заказчиком, сроки этапов, области ответственности заказчика и исполнителя. | 1. Сбор данных из различных источников (внутренние, подрядчики).  2. Анализ возможности их регулярного обновления, актуализации (важно для моделей ушедших в развертывание).  3. Описание данных.  Каждый признак объекта должен иметь ясный алгоритм своего формирования, прозрачное, однозначное определение.  4. Проверка качества данных: согласованность данных из различных источников, целостность непротиворечивость.  Набор таблиц базы данных, с описанием сущностей хранимых в них и связями между ними (ER-модель). | 1. Объединение данных.  2. Очистка данных от выбросов.  3. Генерация производных признаков на основе исходного признакового описания.  Обучающая выборка | 1. Определение методов моделирования.  2. Статический и корреляционный анализ данных.  3. Изучение стабильности признаков.  5. Обучение моделей, отбор признаков, подбор гиперпараметров.  6. Валидация моделей.  7. Анализ результатов, сравнение и селекция моделей полученных различными методами моделирования.  Статистическая модель | 1. Оценка финансового эффекта.  2. Создание отчета о моделировании.  3. Презентация модели заказчику. | 1. Планирование развертывания.  2. Планирование поддержки и мониторинга развернутого решения.  3. Создание финального отчета. |