1. Сбор данных
   1. Что собирать?
   2. Откуда собирать?
   3. Какие есть инструменты для сбора?
   4. Когда остановиться?
   5. Как хранить?
2. Предобработка данных
   1. Общая схема процесса предобработки
      1. Очистка данных
      2. Интеграция данных
      3. Трансформация данных
      4. Преобразование данных
   2. Предобработка в NLP
      1. Токенизация
      2. Удаление стоп-слов
      3. Лемматизация и стемминг
      4. Представления текста в виде вектора
      5. Эмбеддинг
   3. Плюсы и минусы различных методов преобразования данных
3. Выбор модели
   1. Что такое выбор модели
   2. Определение подходящей модели
   3. Методы выбора модели
      1. Вероятностные критерии выбора статистической модели
      2. Методы сэмплирования
4. Обучение модели
   1. Типы обучения
      1. Обучение с учителем
      2. Обучение без учителя
      3. Обучение с подкреплением
   2. Методы оптимизации
      1. Классический градиентный спуск
      2. Метод моментов
      3. Адаптивные методы
      4. Стохастический градиентный спуск
   3. Борьба с переобучением
      1. Суть переобучения
      2. Регуляризация
      3. Использование тестовых выборок
5. Оценка обученной модели
   1. Тестирование без участия пользователей
      1. Сбор и генерация данных для тестов
      2. Оценка качества классификации
      3. BLEU
      4. ML метрики оценки качества
   2. Тестирование с участием пользователей
      1. A/B тестирование
      2. Многорукие бандиты
   3. Что оценивать кроме качества?   
      (скорость работы, робастность и тд)