

Полная программа

# Рекомендательные системы

Recommender Systems

Длительность курса: 88 часов

## Модуль 1. От классического ML к персонализации

### Тема 1

**Знакомство. Вводное занятие. Жизненный цикл моделей**

**Цель занятия**  
познакомиться с программой курса; узнать этапы жизненного цикла моделей персонализации и рекомендации.

**Краткое содержание**  
знакомство; обзор модулей курса; этапы жизненного цикла моделей персонализации и рекомендации.

### Тема 2

**Сегментация пользователей и задача персонализации**

**Цель занятия**  
научиться проводить сегментацию пользователей на основе эвристики; научиться применять кластеризацию для задачи сегментации; научиться проводить сегментацию при помощи логистической регрессии.

**Краткое содержание**  
сегментация пользователей на основе эвристики; K-means и DBSCAN для задачи сегментации; сегментация при помощи логистической регрессии.

### Тема 3

**Look-a-like выделение сегментов пользователей**

**Цель занятия**  
формализовать задачу look-a-like; научиться находить похожих пользователей при помощи моделей машинного обучения.

**Краткое содержание**  
постановка задачи look-a-like выделения сегментов пользователей; поиск похожих пользователей при помощи моделей машинного обучения.

### Тема 4

**Модели Next Best Action**

**Цель занятия**  
формализовать задачу Next Best Action; научиться решать задачу Next Best Action при помощи моделей машинного обучения.

**Краткое содержание**  
постановка задачи Next Best Action; решение задачи Next Best Action при помощи моделей машинного обучения.

### Тема 5

**Uplift - моделирование**

**Цель занятия**  
формализовать задачу Uplift - моделирования; научиться решать задачу Uplift - моделирования при помощи построения одной и нескольких моделей.

**Краткое содержание**  
формализация задачи как набор нескольких моделей; построение в виде одной модели; построение в виде нескольких моделей.

### Домашние задания

**Uplift - моделирование**  
**Цель**  
В этом ДЗ вам нужно реализовать модель, решающую задачу Uplift-моделирования с помощью перекрестной зависимости.

## Модуль 2. Классические методы рекомендаций

### Тема 1

**Введение в рекомендательные системы**

**Цель занятия**  
формализовать задачу рекомендации; изучить типы данных и методы разбиения на обучающую и тестовую выборки; научиться вычислять метрики оценки качества; изучить основные классы рекомендательных моделей.

**Краткое содержание**  
постановка задачи рекомендательной системы; explicit и implicit типы фидбека; разбиение данных на обучающую и тестовую выборки; метрики оценки качества; рекомендательные модели в продакшн; классификация рекомендательных моделей.

### Тема 2

**Эвристические модели. Коллаборативная фильтрация**

**Цель занятия**  
научиться решать проблему холодного старта; изучить эвристические подходы; понять методы коллаборативной фильтрации; научиться применить метод поиска ближайших соседей к задаче рекомендаций.

**Краткое содержание**  
эвристические модели; проблема холодного старта; userKNN, itemKNN.

### Тема 3

**Методы матричной факторизации**

**Цель занятия**  
понять методы матричной факторизации; научиться применять алгоритмы SVD и ALS.

**Краткое содержание**  
алгоритм ALS; алгоритм SVD; применение алгоритмов на практике.

### Тема 4

**Практическое занятие по моделям матричной факторизации**

**Цель занятия**  
научиться применять SVD и ALS - алгоритмы для разных типов данных.

**Краткое содержание**  
модификации SVD и ALS - алгоритмов.

### Домашние задания

**Модели матричной факторизации**  
**Цель**  
В этом задании вам нужно сравнить качество моделей матричной факторизации.

### Тема 5

**A/B - тесты в рекомендательных системах**

**Цель занятия**  
изучить бизнес-метрики; понять методы проведения A/B - тестирования.

**Краткое содержание**  
бизнес-метрики; методика проведения A/B - тестов; планирование и оценка длительности A/B - теста; подведение итогов A/B - теста.

## Модуль 3. Контентные и гибридные методы рекомендаций

### Тема 1

**Контентные методы рекомендаций**

**Цель занятия**  
изучить особенности контентных методов рекомендаций; научиться применять MF и LightFM - модели.

**Краткое содержание**  
факторизационные машины; LightFM.

### Домашние задания

**Использование фичей для улучшения качества рекомендаций**  
**Цель**  
В этом задании вы продолжите работу с LightFM, исследуете её качество как модели коллаборативной фильтрации и при холодном старте.

### Тема 2

**Рекомендательная система на основе текстовых данных**

**Цель занятия**  
научиться строить эмбединги слов на основе предобученного FastText; научиться суммаризировать тексты; научиться строить эмбединги BERT LaBSE; научиться искать похожие на основе эмбедингов.

**Краткое содержание**  
fasttext; LSI; BERT; gensim.

### Тема 3

**Ранжирование при помощи градиентного бустинга**

**Цель занятия**  
изучить особенности задачи ранжирования каталога товаров; научиться применять эвристические подходы к задаче ранжирования; научиться применять модели машинного обучения для ранжирования каталога товаров.

**Краткое содержание**  
эвристические модели; градиентный бустинг.

### Тема 4

**Практическое занятие по рекомендательным системам Двухуровневая модель**

**Цель занятия**  
построить двухуровневую модель.

**Краткое содержание**  
для чего нужна двухуровневая модель; построение двухуровневой модели.

### Домашние задания

**Двухуровневая модель**  
**Цель**  
В этом задании вы продолжите работу с двухуровневой моделью, попробуете добавлять новые признаки, а также модели первого уровня, подберете гиперпараметры. В результате доработок качество модели, рассмотренной на занятии, должно улучшиться.

## Модуль 4. Современные методы рекомендаций

### Тема 1

**Введение в нейросети и PyTorch**

**Цель занятия**  
разобрать как устроена нейронная сеть; научиться обучать нейронную сеть; научиться решать проблему переобучения; познакомиться с фреймворком для глубокого обучения PyTorch.

**Краткое содержание**  
подготовка данных для обучения; dataset, dataloader; порядок обучения нейронной сети; регуляризация, dropout, early stopping.

### Тема 2

**Нейросетевые методы рекомендаций**

**Цель занятия**  
понять особенности нейросетевой коллаборативной фильтрации, sequential, session-based подходы.

**Краткое содержание**  
нейросетевая коллаборативная фильтрация; sequential, session-based подходы.

### Домашние задания

**Нейросетевые модели**  
**Цель**  
В этом задании вам нужно построить модель NCF и исследовать её качество.

### Тема 3

**Sequential - модели**

**Цель занятия**  
познакомиться с механизмом self-attention; рассмотреть архитектуру моделей SasRec и Bert4Rec.

**Краткое содержание**  
механизм self-attention; SasRec; Bert4Rec.

### Тема 4

**Методы теории графов в рекомендательных системах**

**Цель занятия**  
понять, как могут помочь графовые структуры данных для предсказания рекомендаций; изучить практические подходы.

**Краткое содержание**  
разбор методов и алгоритмов; разбор пайплайна решения.

### Тема 5

**Многорукие бандиты для оптимизации A/B - тестирования**

**Цель занятия**  
изучить виды алгоритмов многоруких бандитов; рассмотреть, как обучение с подкреплением используется для A/B - тестирования.

**Краткое содержание**  
Multi-armed bandits; виды алгоритмов многоруких бандитов; применение многоруких бандитов для A/B - тестирования.

### Тема 6

**Многорукие бандиты для задачи рекомендации**

**Цель занятия**  
научиться применять многорукие бандиты в задаче рекомендации.

**Краткое содержание**  
Multi-armed bandits; применение многоруких бандитов для задачи рекомендаций.

## Модуль 5. Рекомендательные системы в продакшн

### Тема 1

**Особенности работы с большими объемами данных**

**Цель занятия**  
изучить особенности работы с большими данными; изучить особенности фреймворков для работы с большими данными.

**Краткое содержание**  
Pandas; Polars; PySpark.

### Тема 2

**Внедрение рекомендательных систем на практике. Микросервисы**

**Цель занятия**  
рассмотреть этапы внедрения рекомендательных моделей в продакшн; изучить особенности построения микросервисной архитектуры.

**Краткое содержание**  
этапы внедрения рекомендательных моделей в продакшн; микросервисы.

### Тема 3

**Применение Docker на практике**

**Цель занятия**  
изучить, для чего нужен Docker; научиться использовать Docker на практике.

**Краткое содержание**  
Docker и его структура; использование Docker контейнеров на практике.

## Модуль 6. Проектная работа

### Тема 1

**Выбор темы и организация проектной работы**

**Цель занятия**  
выбрать и обсудить тему проектной работы; спланировать работу над проектом; ознакомиться с регламентом работы над проектом.

**Краткое содержание**  
правила работы над проектом и специфика проведения итоговой защиты; требования к результату проекта и итоговой документации.

### Домашние задания

**Построение рекомендательной системы для решения задачи**  
**Цель**  
В проектной работе вам нужно:  
- Закрепить и продемонстрировать полученные знания и навыки.  
- Отработать навыки командной работы (не более 2 человек).  
- Подготовить портфолио для работодателя.

### Тема 2

**Консультация по проектам и домашним заданиям**

**Цель занятия**  
получить ответы на вопросы по проекту, ДЗ и по курсу.

**Краткое содержание**  
вопросы по улучшению и оптимизации работы над проектом; затруднения при выполнении ДЗ; вопросы по программе.

### Тема 3

**Защита проектных работ и подведение итогов курса**

**Цель занятия**  
защитить проект и получить рекомендации экспертов.

**Краткое содержание**  
презентация проектов перед комиссией; вопросы и комментарии по проектам.