OTUS

Полная программа

Рекомендательные системы

Recommender Systems

Длительность курса: 16 часов

Модуль 1. Классические методы рекомендаций

Тема 1
Введение в рекомендательные системы. Рекомендательные системы, основанные на правилах

Цель занятия
рассмотреть рекомендательные системы, как их часть бизнес-процессов;
научиться строить поиграть для Библиотеки, построенных на правилах для различного типа сервисов:
новостные порталы;
новостные агрегаторы;
контентные серверы;
форумы;
социальные сети.

Задачи занятия
ABC анализ товаров;
RFM анализ пользователей;
статистический анализ продаж и просмотров контента.

Тема 2
Коллаборативная фильтрация на примере рейтингов фильмов

Цель занятия
разобрать алгоритмы коллаборативной фильтрации;
разобрать алгоритмы KNN;
разобрать метрики;
по темам:
поиманиентности с библиотекой surprise.

Задачи занятия
user-based knn;
item-based knn;
surprise lib.

Домашнее задание

Построение рекомендательной алгоритмики KNN

Цель
В этом ДЗ вы подобьете с литературоведение алгоритмы KNN.

Тема 3
SVD-разложение на примере рейтингов фильмов

Цель занятия
разобрать алгоритм SVD;
разобрать алгоритм SVD++;
построить поиграть рекомендацию алгоритмы SVD++.

Задачи занятия
surprise lib;
svd;
svd++.

Тема 4
SVD-разложение на примере продаж товаров

Цель занятия
рассмотреть обобщение алгоритма SVD в виде классификационной задачи;
научиться добавлять в алгоритм информацию о пользователях;
научиться добавлять в алгоритм информацию о последнем покупках.

Задачи занятия
factorizationmachine;
factorizationmachine на sklearn.

Тема 5
Метрики качества и рекомендательных и A/B-тесты

Цель занятия
научить алгоритмы задач ранжирования;
рассмотреть примеры бизнес метрик;
изучить алгоритмы суперсрезов A/B тестов: оценка количества примеров для тестов;
p-value (значимо примеров потребно 16).

Задачи занятия
расчет метрик:
precision@k;
recall@k;
Average precision at K;
map@K;
ndcg@K;
CTR.

для в продажах, выручка у рекомендации;
работа с алгоритмами метрикам A/B тестов:
выборка примеров A/B тестов и корреляционная пользователей.

Домашнее задание

Построение рекомендательной алгоритмики SVD, SVD++

Цель
В этом ДЗ вы сделаете работу с алгоритмами матричного разложения.

Модуль 2. Контентные методы рекомендаций

Тема 1
Контентные методы рекомендаций

Цель занятия
познакомиться с контентными методами рекомендаций;
рассмотреть способы выдачи рекомендаций для новых пользователей и новых айтемов.

Задачи занятия
обоснованные прибегнем коллаборативной фильтрации;
методы обобщения рекомендаций на основе контента айтемов.

Тема 2
Рекомендательная система на основе текстовых данных

Цель занятия
научится строить embeddings слов на основе предобученного FastText;
научиться суммаризировать тексты;
научиться строить рекомендательные системы для подбора товаров по embeddings;
научиться оценивать для разных рекомендаций по embeddings (сравнение полученных суммаризаций и подбор похожих товаров с суммаризацией embeddings);

Задачи занятия
nltk;
razdel[10];
SentiRuEval[11];
MLP;
DSSM;
pytorch.

Домашнее задание

Реализация рекомендательной системы на основе текста

Цель
В этом ДЗ вам нужно построить вектора текста для целей рекомендательной системы, использует предобученные embeddings FastText для русского языка, или предобученную модель типа BERT для русского языка.

Тема 3
Рекомендательная система на основе анализа изображений

Цель занятия
научиться embeddings изображений;
рассмотреть способы выдачи рекомендаций для новых изображений;
разобрать примеры рекомендательной системы для изображений (сравнение полученных суммаризаций и подбор похожих товаров с суммаризацией embeddings);

Задачи занятия
keras[12];
keras[13];
MLP;
DSSM;
pytorch.

Тема 4
Информационный поиск как рекомендательная система

Цель занятия
изучить задачу информационного поиска;
рассмотреть алгоритмы поиска по первым и изучить формуру ранжирования;
изучить инструмент подходы к поиску;
рассмотреть подходы к ранжированию и поиск/ранжир ведущих рекомендательных платформ.

Задачи занятия
TF-IDF;
OKAPI BM25;
LSI;
LDA;
DSSM.

Модуль 3. Современные методы рекомендаций

Тема 1
Нейронные сети в рекомендательных системах

Цель занятия
познакомиться с библиотекой DJRec;
реализовать на уровне примера модели NeuralCF.

Задачи занятия
pytorch;
learning rate scheduler;
научиться подбирать для дополнительных данных о пользователях, айтемах.

Домашнее задание

Реализация рекомендательной системы на основе нейросетевых алгоритмы

Цель
В этом ДЗ вы реализуете NeuralCF на датасете movielens-1M.

Тема 2
Многорукие бандиты

Цель занятия
реализовывать алгоритмы моделей на основе вероятностного моделирования.

Задачи занятия
реализация работы рекомендательной системы на основе логик клиентбаннеров;
сравнение рекомендательных систем на основе вероятностной системы.

Тема 3
Гибридные рекомендательные системы

Цель занятия
обзор методов смешения систем
mixed;
weighted;
switching;
feature augmented;
feature augmentation.

Задачи занятия
разобор кейсов построения гибридных рекомендательных систем.

Тема 4
Внедрение рекомендательных систем на практике

Цель занятия
обзорный архитектуры продукции внедрения рекомендательных систем;
рассмотреть способы внедрения;
рассмотреть минимальную архитектуру для онлайн-рекомендаций;
научиться работать с фреймворком Tarantool или Redis.

Задачи занятия
aerl - фреймворк Redis;
tarantool - фреймворк;
результаты/ошибки;
принцип каскада моделей;
принцип ансамбля моделей;
хранение моделей в клиентском формате;
работа с clickhouse.

Модуль 4. Персонализация

Тема 1
Сегментация пользователей и задача персонализации

Цель занятия
научиться проводить сегментацию пользователей на основе RFM и кластерий - учебный пример ДЗ - магазин;
научиться кластеризовать матрицы клиент, фичи;
научиться сегментировать клиентов на клиентские данные - делать значимые разницы(p-value);
научиться сегментировать на основе мультиклассовой задачи и обучающих примеров клиентов.

Задачи занятия
K-means;
DBSCAN;
Agglomerative;
автоэнкодер из MLP + кластеризация на выбодщатели.

Домашнее задание

Кластеризация клиентов интернет магазина. Поиск кандидатов в VIP-клиенты

Цель
В этом ДЗ вы реализуете сегментацию клиентов на основе кластеризации. Вам нужно будет выделить VIP-клиентов сегментацию клиентов на основе кластеризации.

Тема 2
Look a-like выделение сегментов пользователей

Цель занятия
сформулировать задачу look a-like, как задачу классификации;
научиться подготовкам решения в виде классификации задачи, подготовкам решения в виде классификации моделей LDA, logreg, catboost, lookalikeforest;
научиться работать с формированием значимой выборки, или обращениям TSFF - рекомендацией от самодеятельности алгоритмы.

Задачи занятия
classificaiton catboost;
classificaiton logreg;
classificaiton lookalikeforest;
classificaiton LDA;
обращениям TSFF - списков моделей.

Тема 3
Модели Next Best Action

Цель занятия
научиться формулировать задачу в виде построения набора классификаторов;
научиться строить набор классификаторов на данных примеров;
научиться оценивать такие модели;
сделать сравнительную постановщика рекомендательных предложений.

Задачи занятия
classificaiton lookalikeforest;
xgboost;
catboost;
алгоритм выбора рекомендации для пользователя.

Тема 4
Uplift моделирование

Цель занятия
рассмотреть формализацию задачи, как набор независимых моделей;
изучить постановкам в виде целей моделей;
изучить постановкам в виде нескольких моделей.

Задачи занятия
построение моделей Uplift (на основе нейросети;
построение моделей Uplift (на основе catboost.

Модуль 5. Проектная работа

Тема 1
Выбор темы и организация проектной работы

Цель занятия
выбрать и обсудить тему проектной работы;
сгенерировать работу над проектом;
скоординировать с работами над проектом.

Задачи занятия
правка работы над проектом и специфика проведения итоговой аттестации;
требования к результату проекта и итоговой документации.

Домашнее задание

Проектная работа. Построение гибридной рекомендательной системы

Цель
В проектной работе, вам нужно реализовать рекомендательную систему, как гибрид из нескольких рекомендательных алгоритмов на языке python.

Тема 2
Консультация по проектам и домашним заданиям

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

Задачи занятия
вопросы по алгоритмам и оптимизации работы над проектом;
затруднения при выполнении ДЗ;
вопросы по программам.

Тема 3
Защита проектных работ и подведение итогов курса

Цель занятия
защитить проект и получить рекомендации экспертов.

Задачи занятия
презентация проектов перед комиссией;
вопросы и комментарии по проектам.