Курс "Рекомендательные системы", ОП ПМИ
3 курс

Общая информация по курсу

Аннотация	*Курс нацелен на формирование у слушателя достаточных знаний, навыков и компетенций для построения рекомендательных систем. Курс является прикладным и охватывает все продвинутые темы, необходимые для успешного применения как в индустрии, так и в академических исследованиях. Слушатели курса научатся постановке и имплементации рекомендательных ML моделей на Python от простой коллаборативной фильтрации до современных нейронных сетей.*
Цели освоения курса	1. *Знать основные подходы рекомендательных систем и алгоритмов к решению задачи ранжирования и рекомендаций пользователям.* 2. *Уметь правильно определять и формализовывать задачу для построения рекомендательной системы.* 3. *Определять и имплементировать функции ранжирования и метрики качества оценки алгоритмов.* 4. *Разрабатывать модели машинного обучения для рекомендательных систем разной сложности на языке программирования Python.* 5. *Объяснять и интерпретировать предсказания на локальном и глобальном уровне моделей, использовать их для отладки ошибок и аномалий.* 6. *Разрабатывать базовые API сервисы для рекомендательных моделей с помощью docker.*
Пререквизиты курса (необязательно)	1. *Линейная алгебра* 2. *Теория вероятностей* 3. *Программирование на Python* 4. *Алгоритмы и структуры данных* 5. *Машинное обучение* 6. *Глубинное обучение*
Технические требования	*Python > 3.11* *Jupyter Notebook*

Понедельный план курса

№	Тема лекции	Тема семинара
1	Введение в рекомендательные системы	Обзор прикладных задач для внедрения рекомендательных систем. Формализация задачи ранжирования выдачи поиска и персональных рекомендаций. Подбор подходящего таргета для модели с учетом бизнес-метрики. Implicit и explicit тип данных. Виды функций ранжирования и метрик оценки качества ранжирования. Типовые проблемы рекомендательных систем: popularity bias, cold start problem, feedback loop, filter bubble.
2	Базовые подходы. Item-2-item рекомендации	Неперсональные рекомендации и бенчмарки на статистиках. Подходы на основе user-based и item-based метрик близости. Прикладные задачи и подходы для item-item рекомендаций. Item KNN для RecSys (например, похожие и сопутствующие объекты). Линейные модели на примере EASE, SLIM.
3	Матричная факторизация	Матричные разложения на примере SVD. Проблемы разреженности и высокой размерности данных. Тензорные разложения.
4	Коллаборативная фильтрация	Имплементация implicit ALS. Техники препроцессинга матрицы интеракций. Нейросетевая коллаборативная фильтрация Neural CF.
5	Контентные подходы	Способы добавления user и item признаков в модели. Разбор моделей LightFM и DSSM.
6	Гибридные модели	Таксономия гибридных подходов. Архитектура каскадной (двухуровневой) модели. Типовые ошибки и особенности разбиения датасета, кросс-валидации, подбора параметров пайплайна. Градиентный бустинг как end-to-end рекомендательная модель и как ранкер второго уровня. Нейросетевой ранжирование как ранкер.
7	Нейросетевые модели для задачи next-item рекомендаций	Постановка задачи рекомендаций на последовательных данных. GRU4Rec, SASRec, BERT4Rec.
8	Контекстные рекомендации	Контекстные модели. Time-aware, time-dependent рекомендательные подходы. Temporal reasoning. Техники учета времени в моделях.
9	Нейросетевые модели для задачи NBR (next-basket recommendation)	Задача предсказания следующей корзины. Алгоритмы TIFU KNN, TAIW и другие алгоритмы.

10	Автоэнкодеры и вариационные автоэнкодеры	Особенности использования автоэнкодеров и вариационных автоэнкодеров для рекомендаций. Имплементация Rec-VAE, Multi-VAE, Multi-DAE.
11	Графовые нейросетевые модели	Графовые рекомендательные системы GCN, GFCF, LightGCN. Индуктивные модели на примере PinSage. Проблема out-of-sample пользователей и объектов
12	Интерпретируемость и объяснимость моделей	Model-intrinsic и model-agnostic подходы интерпретации рекомендательных систем. Векторы Шепли. Интерпретация весов внутри модели (пример iALS), через механизм внимания. Обзор нейросетевых подходов для рекомендаций с объяснениями
13	Uplift и экономические рекомендательные системы	Разбор прикладных метрик и подходов для задачи получения uplift от показа рекомендаций вместо задачи предугадывания действий.
14	Multi-task и cross-domain рекомендательные модели	Рекомендательные системы с обучением на нескольких источниках данных и нескольких таргетах. Attribute, type и item уровни обучения. Content-based, embedding-based, rating pattern based transfer подходы. Single, dual и multi-task таргеты в CDR моделях. Contrastive learning модели.
15	RL для рекомендательных систем	Offline и online RL-based подходы для рекомендательных систем. TD3, TD3+BC.
16	Контентные рекомендации на мультимодальных данных	Рекомендательные системы на основе текстов, изображений и видео. Transformed-based модели.
17	A/B тестирование рекомендательной системы	Онлайн и оффлайн метрики, поиск proxy метрик. Оценка вспомогательных метрик качества рекомендаций (novelty, diversity, serendipity, coverage). Дизайн мультивыборочного тестирования и оценки эффекта самого пользовательского интерфейса. Использование многоруких бандитов как альтернативы A/B тестов.
18	Large scale RecSys	Особенности индустриальной разработки рекомендательных систем. RecSys ML System Design. Типовые архитектуры рекомендательных систем. Имплементация моделей для высоконагруженных систем.
19	Написание сервиса рекомендательной	Разбор и имплементация простой архитектуры сервиса рекомендательной системы. Использование Docker, FAST API,

	системы	Inference server (Triton), Redis для поднятия сервиса. Мониторинг стабильности и деградации модели.
20	Прикладные аспекты и тренды в рекомендательных системах	Дополнительные техники и подходы, используемых на практике. Индустриальные и академические тренды в рекомендательных системах. Адаптация LLM для построения рекомендательной системы и оценки их качества.

Практика и контроль

Форма контроля	Содержание	Когда выдаётся/проходит?
Домашнее задание 1	*Выдается после семинара в формате Jupyter Notebook на имплементацию функций ранжирования и метрик оценки качества. Содержит 5 заданий, каждое из которых оценивается в 0, 1 или 2 балла (т.е. максимум 10 баллов).*	*Выдается после семинара №1*
Домашнее задание 2	*Выдается после семинара в формате Jupyter Notebook на имплементацию iALS. Содержит несколько заданий, каждое из которых оценивается в 0, 1 или 2 балла (максимальная оценка 10 баллов).*	*Выдается после семинара №4*
Домашнее здание 3	*Выдается после семинара в формате Jupyter Notebook на имплементацию двухуровневого пайплайна. Содержит несколько заданий, каждое из которых оценивается в 0, 1 или 2 балла (максимальная оценка 10 баллов).*	*Выдается после семинара №6*
Домашнее задание 4	*Выдается после семинара в формате Jupyter Notebook на имплементацию нейросетевой модели на последовательных данных. Содержит несколько заданий, каждое из которых оценивается в 0, 1 или 2 балла (максимальная оценка 10 баллов).*	*Выдается после семинара №8*

Домашнее задание 5	*Участие в открытом соревновании по рекомендательным системам на онлайн платформе (Kaggle). Градация баллов по 10-балльной шкале в соответствии с финальным лидербордом, выданная одновременно с заданием.*	*Выдается после семинара №11*
Тест	*Выставляется средняя арифметическая оценка за 15 тестов, которые проводятся на базе Google Forms/Anytask. Каждый тест содержит от 4 до 10 вопросов на пройденную тему последних лекции и семинаре в формате открытых вопросов либо с выбором ответов. Максимальная оценка 10 баллов.*	*Выдается после каждого семинара*
Экзамен	*Устный, дистанционный экзамен на онлайн платформе Zoom. Разрешается пользоваться материалами. Экзамен неблокирующий. Продолжительность - 20 минут на студента. Максимальная оценка 10 баллов на основе оценки правильности ответов преподавателем.*	В конце курса
Пересдача	*Выставляется кумулятивная оценка. Если по окончанию курса балл недостаточный на удовлетворительно (4 балла по 10-балльной шкале), проводится пересдача всех форм контроля.*	

Формула оценки и пересдачи

Пример формулы:
*Итог = Округление(0.3 * ДЗ + 0.3 * Т + 0.4 * Э),*
где ДЗ — средняя оценка за все домашние задания, Т — средняя оценка за все тесты, Э — оценка за экзамен.
Округление арифметическое.

Описания автоматов:
*Определим накопленную оценку как Накоп = Округление((0.3 * ДЗ + 0.3 * Т + 0.4 * Э) / 0.7). Если Накоп >=8 и среднее арифметическое за тесты не ниже 8 баллов или выше, то студент может получить Накоп в качестве итоговой оценки, не приходя*

на экзамен. Значения порогов могут быть пересмотрены с учетом фактической общей успеваемости потока.
Пересдача:
Первая пересдача проводится в формате сдачи всех форм контроля, ни одна из которых не является блокирующей. Формула итоговой оценки не меняется. Вторая пересдача (с комиссией) проводится в формате, аналогичном экзамену. Формула оценки не меняется за одним исключением: если за вторую пересдачу студент получил оценку 7 и выше, то гарантируется, что итоговая оценка будет не ниже 4 баллов.

Литература

Рекомендуемая основная литература

1. Recommender Systems Handbook. Francesco Ricci et al. (Eds.), Springer, 2022 – Режим доступа: https://link.springer.com/book/10.1007/978-1-0716-2197-4

Рекомендуемая дополнительная литература

1. Charu C. Aggarwal. Recommender Systems: The Textbook, Springer, April 2016 – Режим доступа: https://www.springer.com/gp/book/9783319296579