

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«24» июня 2025 г.
Протокол № 2

**Рабочая программа дисциплины (модуля)
«Рекомендательные системы»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Машинное обучение

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2025

**Москва
2025**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение	8
6. Материально-техническое обеспечение	8
7. Методические и оценочные материалы	10

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Рекомендательные системы» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по специальности 02.04.01 Математика и компьютерные науки, профиль Машинное обучение, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Рекомендательные системы» позволяет овладеть современными методами персонализации цифровых сервисов, что является ключевым фактором повышения качества пользовательского опыта и эффективности ИТ-продуктов в различных отраслях.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Машинное обучение и входит в обязательную часть Блока 1.

Дисциплина (модуль) изучается на 2 курсе в 4 семестре, доступна для прохождения при условии успешного завершения дисциплины (модуля) «Deep Learning (Глубокое обучение)».

Цель изучения дисциплины (модуля): формирование системных знаний, практических умений и навыков разработки, оценки и внедрения современных рекомендательных систем на основе классических и продвинутых методов машинного обучения и глубокого обучения.

Задачи изучения дисциплины (модуля):

- изучить основные подходы к построению рекомендательных систем, классификацию функций ранжирования и современные модели рекомендаций;
- освоить методы формализации задачи рекомендаций, выбора и расчёта метрик качества ранжирования;
- научиться разрабатывать и реализовывать рекомендательные модели различной сложности на Python с использованием PyTorch и специализированных фреймворков (RecBole, RecTools, RePlay);
- сформировать навыки интерпретации и анализа предсказаний моделей, а также выявления и устранения ошибок и аномалий;
- освоить практики внедрения рекомендательных систем, включая разработку API-сервисов (Docker) и проведение A/B-тестирования.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- основные подходы рекомендательных систем и алгоритмов к решению задачи ранжирования и рекомендаций пользователям;
- основные метрики для задачи ранжирования и рекомендаций;
- классификацию функций ранжирования и основных представителей каждого класса;
- продвинутые модели для построения рекомендаций.

уметь:

- правильно определять и формализовывать задачу для построения рекомендательной системы;
- определять и имплементировать функции ранжирования и метрики качества оценки алгоритмов;
- разрабатывать модели машинного обучения для рекомендательных систем разной сложности на языке программирования Python;
- объяснять и интерпретировать предсказания на локальном и глобальном уровне моделей, использовать их для отладки ошибок и аномалий;

- разрабатывать базовые API сервисы для рекомендательных моделей с помощью Docker.

владеть:

- основными рекомендательными фреймворками: RecBole, RecTools, RePlay;
- навыком написания глубоких нейронных сетей на PyTorch для рекомендательных систем;
- метриками для задачи ранжирования;
- инструментами для проведения A/B тестирования.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
ОПК-3.	Способен самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства	ОПК-3.1.	Знает основные принципы программирования, архитектуры программного обеспечения и современные языки программирования, а также особенности отечественных информационных технологий и сетевых ресурсов
		ОПК-3.2.	Умеет разрабатывать прикладные программные средства, используя современные инструменты и технологии, а также интегрировать их с сетевыми ресурсами для решения конкретных задач
		ОПК-3.3.	Имеет практический опыт разработки программных средств, используемых при построении математических моделей в естественных науках
ПК-2.	Способен математически корректно ставить естественнонаучные и прикладные задачи	ПК-2.1.	Знает основные методы и подходы к математическому моделированию, а также теоретические основы естественных и прикладных наук, необходимые для корректной формулировки задач
		ПК-2.2.	Умеет анализировать практические ситуации и формулировать на их основе математические модели, включая выбор адекватных методов решения и формулировку условий задачи
		ПК-2.3.	Имеет практический опыт в разработке и решении математических задач в рамках проектов или научных исследований, где были успешно поставлены и решены естественнонаучные и прикладные задачи
ПК-5.	Способен передавать результат решенных прикладных задач в виде	ПК-5.1.	Знает основные методы и подходы к формулированию рекомендаций на основе

	конкретных рекомендаций, выраженных в терминах области машинного обучения		результатов решения прикладных задач, а также термины и концепции, специфичные для области машинного обучения
		ПК-5.2.	Умеет анализировать результаты решенных задач и формулировать четкие, конкретные рекомендации, адаптируя их к требованиям и ожиданиям целевой аудитории
		ПК-5.3.	Имеет практический опыт в разработке и представлении рекомендаций на основе анализа прикладных задач, включая участие в проектах, где результаты были успешно применены и оценены в контексте области машинного обучения

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостоя тельная работа	
Лекции	Семинары (практичес кие занятия)					
1	Введение в рекомендательные системы	8	8		33	Домашние задания, Тесты
2	Deep Learning в рекомендательных системах	6	6		25	Домашние задания, Тесты
3	Продвинутый Deep Learning	6	6		25	Домашние задания, Тесты
4	Важные практические темы в области рекомендаций	8	8		33	Домашние задания, Тесты
5	Тренды направления	2	2		10	Домашние задания, Тесты
	<i>Экзамен</i>			4		
	Итого:	30	30	4	126	
	<i>Объем дисциплины (модуля) (в ак. ч.)</i>	190				
	<i>Объем дисциплины (модуля) (в зач. ед.)</i>	5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Введение в рекомендательные системы	Введение в рекомендательные системы Базовые подходы. Item-2-item рекомендации Матричная факторизация. Коллаборативная фильтрация Двухуровневый пайплайн
2	Deep Learning в рекомендательных системах	Глубокие нейрорекомендации для topn recommendations. Нейросетевые модели для задачи next-item рекомендаций Нейросетевые ранкеры
3	Продвинутый Deep Learning	Контентные подходы. Рекомендации на мультимодальных данных RL для рекомендательных систем Графовые нейронные сети в рекомендациях
4	Важные практические темы в области рекомендаций	Uplift и экономические рекомендательные системы A/B тестирование рекомендательной системы Распространенные боли в рекомендательных системах Написание сервиса рекомендательной системы
5	Тренды направления	Прикладные аспекты и тренды в рекомендательных системах

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Малов, Д. А. Глубокое обучение и анализ данных. Практическое руководство : практическое руководство / Д. А. Малов. - Санкт-Петербург : БХВ-Петербург, 2023. - 272 с. - ISBN 978-5-9775-1172-8. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2123365>.

2. Рабчевский, А. Н. Синтетические данные и развитие нейросетевых технологий : учебник для вузов / А. Н. Рабчевский. — Москва : Издательство Юрайт, 2025. — 187 с. — (Высшее образование). — ISBN 978-5-534-17716-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/568661>.

Дополнительная литература:

1. Меджедович, Д. Алгоритмы и структуры для массивных наборов данных : практическое руководство / Д. Меджедович, Э. Тахирович ; пер. с англ. А. В. Логунова. – Москва : ДМК Пресс, 2024. - 342 с. – ISBN 978-5-93700-250-1. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2205044>.

2. Равичандиран, С. Глубокое обучение с подкреплением на Python. OpenAI Gym и TensorFlow для профи : практическое руководство / С. Равичандиран. - Санкт-Петербург : Питер, 2020. - 320 с. - (Серия «Библиотека программиста»). - ISBN 978-5-4461-1251-7. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1756109>.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		

КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Рекомендательные системы» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, тесты, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар — это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники

информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Тест – особая форма проверки знаний. Проводится после освоения одной или нескольких тем и свидетельствует о качестве понимания основных понятий изучаемого материала. Тестовые задания составлены к ключевым понятиям, основным разделам, важным терминологическим категориям изучаемой дисциплины (модуля).

Для подготовки к тесту необходимо знать терминологический аппарат дисциплины (модуля), понимать смысл научных категорий и уметь их использовать в профессиональной лексике. Владение понятийным аппаратом, включённым в тестовые задания, позволяет преподавателю быстро проверить уровень понимания студентами важных методологических категорий.

Бонусные баллы – это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Рекомендательные системы»

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме *экзамена*, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования,
9	Отлично	
8	Отлично	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	
5	Удовлетворительно	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Рекомендательные системы» оценивается следующим образом:

Активность	Вес	Количество	Описание
Накопительная оценка			
Домашние задания	70%	6	Набор задач по темам недели
Тесты		15	Письменная работа с набором задач, которые нужно решить за ограниченное время
Промежуточная аттестация			
Экзамен	30%	1	Мок-интервью — имитация в формате реального собеседования по ML System Design рекомендательных систем

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

Итоговая оценка рассчитывается по накопительной по формуле: $\langle 0,75 \times \text{среднее за домашние задания} + 0,25 \times \text{среднее за тесты} \rangle$.

В случае, если студент хочет улучшить оценку, итоговая оценка по дисциплине (модулю) выставляется по формуле: « $0,7 \times$ накопительная оценка ($0,75 \times$ среднее за домашние задания + $0,25 \times$ среднее за тесты) + $0,3 \times$ экзамен».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1.

1. Реализовать базовую модель Item-2-Item рекомендаций на основе косинусного сходства (или Jaccard) для неявных данных.
2. Реализовать модель коллаборативной фильтрации на основе матричной факторизации (ALS или SGD).
3. Сравнить модели по метрикам ранжирования (Precision@K, Recall@K, MAP@K, NDCG@K).
4. Спроектировать и описать двухуровневый пайплайн (candidate generation + ranking) для выбранного датасета.
5. Подготовить краткий отчёт с анализом результатов и выводами.

Домашнее задание 2.

1. Реализовать нейросетевую модель для top-N рекомендаций (например, Neural Collaborative Filtering).
2. Построить модель для next-item рекомендаций (GRU4Rec, SASRec или аналог) и сравнить её с базовой MF.
3. Реализовать простой нейросетевой ранкер (MLP или Wide&Deep) для переранжирования кандидатов.
4. Провести сравнительный анализ моделей по качеству и времени обучения.
5. В отчёте описать архитектуры, гиперпараметры и интерпретацию результатов.

Домашнее задание 3.

1. Реализовать контентную рекомендательную модель с использованием текстовых или мультимодальных признаков (например, эмбединги изображений или описаний).
2. Построить упрощённую RL-среду (симулятор пользователя) и реализовать базовый алгоритм (например, ϵ -greedy или Policy Gradient) для оптимизации долгосрочной награды.
3. Реализовать графовую модель (LightGCN или аналог) для рекомендаций на пользователь-товарном графе.
4. Сравнить подходы по метрикам и описать сценарии их применимости.

Домашнее задание 4.

1. Рассчитать uplift-метрику на синтетических или реальных данных и проанализировать влияние рекомендаций на целевую метрику.
2. Спроектировать A/B-тест для рекомендательной системы: определить гипотезу, метрики, размер выборки и критерии останова.
3. Проанализировать типовые проблемы (cold start, popularity bias, data leakage) на выбранном датасете и предложить способы их решения.
4. Реализовать REST API сервис для одной из моделей (FastAPI + Docker), подготовить Dockerfile и инструкцию по запуску.

Домашнее задание 5.

1. Подготовить аналитический обзор (5–7 страниц или презентацию) по одному из современных трендов в рекомендательных системах (LLM в рекомендациях, генеративные рекомендации, causal recommender systems, privacy-preserving рекомендации и др.).
2. Описать практический кейс внедрения выбранного подхода в индустрии.
3. Проанализировать ограничения и потенциальные направления дальнейших исследований.

Примерные задания по тестам

Тест 1.

1. Где наибольший процент рекомендательного трафика?
 - a) Amazon
 - b) Google Play
 - c) Pinterest
 - d) NetflixОтвет: d.
2. В 1994г GroupLens предложили систему на базе:
 - a) Контентного i2i
 - b) Коллаборативного i2i
 - c) Контентного u2u
 - d) Коллаборативного u2uОтвет: d.
3. Почему решение участников The Netflix Prize не внедрили?
 - a) Большие ансамбли моделей
 - b) Смена парадигмы потребления
 - c) Смена управления компании
 - d) Модели успели устаретьОтвет: a, b.
4. Рекомендательные алгоритмы оценивают через:
 - a) Пользовательские опросы
 - b) Онлайн эксперименты
 - c) Оффлайн моделирование
 - d) ТендерыОтвет: a, b, c.
5. Retention пользователей проще всего замерить в
 - a) Коротком онлайн эксперименте
 - b) Длинном онлайн эксперименте
 - c) Коротком оффлайн моделировании
 - d) Длинном оффлайн моделированииОтвет: b.
6. К неявному (implicit) фидбеку относятся:
 - a) время просмотра
 - b) лайк
 - c) покупка
 - d) отсутствие фидбекаОтвет: a, c, d.
7. Наиболее верный вариант трейн/тест сплита?
 - a) По айтемам
 - b) По юзерам
 - c) По рейтингам

d) По времени

Ответ: d.

8. Diversity, в отличие от serendipity, легко хакнуть при помощи:

- a) Внешних данных
- b) Рандомизации
- c) Ансамбля моделей
- d) Удаления старых айтемов из каталога.

Ответ: b.

9. Какой вид «популярности» устойчив и к эффекту Матфея, и к недостаточной статистике:

- a) Сумма целевых событий (прим. кликов)
- b) Сумма событий за окно (прим. последний месяц)
- c) Конверсия (прим. клики/показы)
- d) Доверительный интервал конверсии (прим. Вилсон)

Ответ: d.

10. Что лучше использовать на пользователе с 1 лайком и новых айтемах без статистики взаимодействий:

- a) Коллаборативный i2i
- b) Контентный i2i
- c) Коллаборативный u2u
- d) SLIM

Ответ: b.

Тест 2.

1. К моделям на графах относятся:

- a) PinSage
- b) TwHIN
- c) LightGCN
- d) SASRec

Ответ: a, b, c.

2. Для применения SVD ($\hat{R} = U_d \Sigma_d V_d^T$) на новом пользователе можно:

- a) Обучить разложение заново
- b) Умножить сырой вектор пользователя на $V_d V_d^T$
- c) Умножить сырой вектор пользователя на $\Sigma_d V_d^T$
- d) Поделить сырой вектор пользователя на $\Sigma_d V_d^T$

Ответ: a, b.

3. Implicit ALS отличается от Explicit:

- a) Наличием кликов в таргете
- b) «Регуляризацией» на отсутствие взаимодействия
- c) Неявным выводом решения
- d) Наличием лайков в таргете

Ответ: b.

4. SASRec отличается от BERT4Rec

- a) Каузальной матрицей внимания
- b) Отсутствием задачи MLM
- c) Сдвигом целевой последовательности на 1 позицию
- d) Способностью работать с последовательностями

Ответ: a, b, c.

5. На этапе ранжирования лучше всего работают

- a) Explicit модели

- b) Implicit модели
- c) Эвристики (свежее, подписки)
- d) SLIM

Ответ: а.

6. На этапе кандидатогенерации лучше всего работают

- a) Explicit модели
- b) Implicit модели
- c) Эвристики (свежее, подписки)
- d) SLIM

Ответ: b, c, d.

7. К моделям нейросетевого ранжирования относятся

- a) Wide&Deep
- b) DCN-V2
- c) HiFormer
- d) SA-TF

Ответ: a, b, c.

8. Интерпретируемость (Interpretability)

- a) Отвечает на вопрос «почему модель дала такой ответ?»
- b) Отвечает на вопрос «что в признаках указывает на такой ответ?»
- c) Более полезна разработчику
- d) Более полезна потребителю

Ответ: a, c.

9. К смещениям в рекомендательных системах относят:

- a) Popularity bias
- b) Position bias
- c) Selection bias
- d) User/Item bias

Ответ: a, b, c.

10. RL в RecSys позволяет

- a) Работать с не IID наблюдениями
- b) Учитывать долгосрочные цели
- c) Жадно рекомендовать айтемы
- d) Построить бейзлайн ранжирования

Ответ: a, b.

Примерное описание и критерии оценивания к мок-интервью

Описание: Мок-интервью — имитация реального технического собеседования по ML System Design с фокусом на рекомендательные системы. Цель — проверить умение студента проектировать, анализировать и обосновывать архитектурные решения в области рекомендательных систем, а также демонстрировать глубокое понимание ключевых концепций машинного обучения и инженерных аспектов.

В ходе мок-интервью студенту будет предложено спроектировать рекомендательную систему для заданного сценария, ответить на вопросы интервьюера, связанные с выбором алгоритмов, обработкой данных, масштабируемостью, объяснимостью, борьбой с смещениями и практическими ограничениями.

Критерии оценивания:

Критерий	Описание	Максимальный балл
Понимание задачи и требований	Способность правильно интерпретировать бизнес-цели и требования системы.	2.0
Архитектура системы	Логичность и полнота предложенной архитектуры, учёт масштабируемости и отказоустойчивости.	2.0
Выбор и обоснование алгоритмов	Обоснование выбора методов рекомендательных систем (матричная факторизация, нейросети и т.д.)	2.0
Обработка данных и качество	Рассмотрение аспектов подготовки данных, борьбы с пропусками и смещениями.	1.5
Объяснимость и интерпретируемость	Умение интегрировать методы объяснимости в систему для повышения доверия пользователей.	1.0
Практические вопросы и ограничения	Учет ограничений вычислительных ресурсов, времени отклика, безопасности и приватности данных.	1.0
Коммуникация и аргументация	Четкость изложения, структурированность ответов и умение вести диалог.	0.5

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Какой способ разделения датасета на трейн/тест в рекомендательных системах наиболее верный? 1. По айтемам 2. По времени 3. По рейтингам 4. По юзерам	2	ОПК-3
2.	Как называется метрика качества рекомендательной системы, определяющая долю рекомендательных списков, для которых среди верхних K объектов присутствует хотя бы один релевантный элемент?	HitRate / HitRate@K / HitRate at K / hit rate / hit Rate / Hit rate / Hit Rate	ОПК-3
3.	Как называется тип откликов в рекомендательных системах, при котором пользователь напрямую выражает своё отношение к объекту (например, ставит рейтинг или лайк)?	Explicit / Явное / Явное взаимодействие / явное / явное взаимодействие / explicit / Явный / явный	ПК-2
4.	Как называется метод матричной факторизации для коллаборативной фильтрации, при котором векторные представления пользователей и объектов оптимизируются поочередно?	ALS / Alternating Least Squares / iALS / implicit Alternating Least Squares	ПК-5

5.	Как называется метод в рекомендательных системах, основанный на вычислении сходства между пользователями или элементами?	подход на базе подобию	ОПК-3
6.	Укажите название алгоритма, который разлагает матрицу взаимодействий пользователей и элементов на латентные факторы.	матричная факторизация	ОПК-3
7.	Укажите число основных компонентов, часто используемых в анализе данных для снижения размерности (например, в PCA).	2/ Два/ два	ПК-2
8.	Как называется метрика оценки качества рекомендаций, измеряющая долю правильно предсказанных релевантных элементов?	precision	ПК-2
9.	Как называется подход к улучшению рекомендаций, использующий обратную связь от пользователей для обучения модели?	обучение с подкреплением	ПК-5
10.	Укажите название метода, повышающего интерпретируемость рекомендаций путем объяснения причин выбора элементов.	объяснимость	ПК-5
11.	Как называется тип моделей в рекомендательных системах, использующий графы для представления связей между пользователями и элементами?	модели на графах	ОПК-3
12.	Укажите название подхода, который включает последовательности взаимодействий для предсказания рекомендаций.	модели на последовательностях	ОПК-3
13.	Укажите название математической модели, описывающей зависимости между несколькими временными рядами в рекомендательных системах.	векторная авторегрессия	ПК-2
14.	Как называется тип откликов в рекомендательных системах, при котором пользователь напрямую выражает своё отношение к объекту (например, ставит рейтинг или лайк)?	явные отклики	ПК-2
15.	Как называется проблема в рекомендательных системах, связанная с предпочтением популярных элементов над менее известными?	смещение популярности	ПК-5
16.	Укажите название технологии, используемой в промышленных системах для обработки больших объемов данных в реальном времени.	поточковая обработка	ПК-5
17.	Как называется техника, улучшающая матричную факторизацию путем добавления дополнительных факторов?	продвинутая факторизация	ОПК-3
18.	Как называется аспект рекомендательных систем, учитывающий специфику домена, например, в e-commerce или стриминге?	доменные специфики	ПК-5