

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

УТВЕРЖДАЮ
Директор по образовательной
деятельности

Минин
«7» *сентября* 2023 С.Т. Князев



РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля	Модуль
1164093	Рекомендательные системы

Екатеринбург

Перечень сведений о рабочей программе модуля	Учетные данные
Образовательная программа Прикладной искусственный интеллект	Код ОП 09.03.03
Направление подготовки Прикладная информатика	Код направления и уровня подготовки 09.03.03

Программа модуля составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Медведев Максим Александрович	кандидат экономических наук, доцент	Доцент	Базовая кафедра «Аналитика больших данных и методы видеоанализа»
2	Балунгу Даниель		ассистент	Базовая кафедра «Аналитика больших данных и методы видеоанализа»

1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Рекомендательные системы

1.1. Аннотация содержания модуля

Модуль описывает принципы проектирования систем рекомендаций и их ранжирования на основе признаков входных данных, как коллаборативных, так и на основе контента. Рассматриваются как основные базовые подходы к рекомендациям, например, на основе матричного разложения и др., так и подходы на основе машинного обучения. Модуль позволяет освоить популярные алгоритмы построения рекомендательных систем, научиться учитывать персонализированную и не персонализированную, а также внешнюю информацию, научиться оценивать и сравнивать метрики рекомендательных систем в тех или иных прикладных задачах.

1.2. Структура и объем модуля

Таблица 1

№	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах
1	Рекомендательные системы	3
ИТОГО по модулю:		3

1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	Не предусмотрено
Постреквизиты и кореквизиты модуля	Не предусмотрено

1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2

Перечень дисциплин модуля	Код и наименование компетенции	Индикаторы достижения компетенции	Планируемые результаты обучения
1	2	3	4
Рекомендательные системы	ПК-2. Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта	ПК-2.1. Разрабатывает приложения систем искусственного интеллекта	ПК-2.1. 3-1. Знает современные языки программирования, библиотеки и программные платформы для функционального, логического, объектно-ориентированного программирования приложений систем искусственного интеллекта (Python, R, C++, C#)

			ПК-2.1. У-1. Умеет разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#)
	ПК-8. Способен создавать и внедрять одну или несколько сквозных цифровых субтехнологий искусственного интеллекта	ПК-8.1. Участвует в реализации проектов в области сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»	ПК-8.1. З-1. Знает фундаментальные правила построения рекомендательных систем и систем поддержки принятия решений, основанных на интеллектуальных принципах, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений» ПК-8.1. У-1. Умеет применять методы и подходы к планированию и реализации проектов по созданию и поддержке системы искусственного интеллекта на основе сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»

1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной форме.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Рекомендательные системы

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Медведев Максим Александрович	кандидат экономических наук, доцент	Доцент	Базовая кафедра «Аналитика больших данных и методы видеоанализа»
2	Даниель Балунгу		ассистент	Базовая кафедра «Аналитика больших данных и методы видеоанализа»

1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Медведев Максим Александрович, доцент, Базовая кафедра «Аналитика больших данных и методы видеоанализа»
- Балунгу Даниель, ассистент, Базовая кафедра «Аналитика больших данных и методы видеоанализа»

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Базовый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания; Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
1	Введение в рекомендательные системы	Важность и применение рекомендательных систем. Типы рекомендательных систем. Данные о взаимодействии пользователя с элементом и их актуальность
2	Коллаборативная (совместная) фильтрация	Подходы, основанные на памяти (основанные на пользователях и элементах). Подходы, основанные на моделях (матричная факторизация, модели скрытых факторов)
3	Фильтрация на основе контента	Извлечение и представление признаков. Показатели сходства для рекомендаций на основе контента
4	Гибридные рекомендательные системы	Сочетание совместной фильтрации и фильтрации на основе контента
5	Контекстно-зависимые рекомендательные системы	Включение контекстуальной информации (время, местоположение и т.д.) в рекомендации
6	NLP и рекомендательные системы	Модели для обработки естественного языка. Применение NLP в анализе отзывов и контента. NLP в персонализированных рекомендациях. Применение

		алгоритмов машинного обучения в NLP. Примеры алгоритмов машинного обучения в NLP на языке Python
7	Оценка рекомендательных систем	Показатели оценки (RMSE, точность, отзыв и т.д.). Автономные и онлайн-методы оценки

1.3. Направление, виды воспитательной деятельности и используемые технологии

Таблица 1.2

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения
Профессиональное воспитание	профориентационная деятельность	Технология формирования уверенности и готовности к самостоятельной успешной профессиональной деятельности Технология самостоятельной работы	ПК-2.1. Разрабатывает приложения систем искусственного интеллекта	ПК-2.1. У-1. Умеет разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#)

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации.

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Рекомендательные системы

Электронные ресурсы (издания)

1. Фальк, К. Рекомендательные системы на практике : практическое руководство / К. Фальк ; пер. с англ. Д. М. Павлова. - Москва : ДМК Пресс, 2020. - 448 с. - ISBN 978-5-97060-774-9. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1210627>.
2. Сорокин, А. Б. Рекомендательные системы: анамнестические и модельные методы : учебно-методическое пособие / А. Б. Сорокин, Л. М. Железняк, Р. Э. Семенов. — Москва :

РТУ МИРЭА, 2022. — 65 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/265739>.

3. Ганегедара, Т. Обработка естественного языка с TensorFlow : практическое пособие / Т. Ганегедара ; пер. с англ. В. С. Яценкова. - Москва : ДМК Пресс, 2020. - 382 с. - ISBN 978-5-97060-756-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1908565>.

Печатные издания

1. Мохов А.И., Кислинский В.Г., Алексейчук А.С. Разработка рекомендательной системы на основе сессий с использованием многоуровневой системы отбора кандидатов // Труды МАИ. 2022. № 126. DOI: 10.34759/trd-2022-126-20
2. Хобсон, Лейн. Обработка естественного языка в действии // Лейн Хобсон, Хапке Ханнес, Ховард Коул ; [перевели с английского И. Пальти, С. Черников]. - Санкт-Петербург [и др.] : Питер, 2020. - 575 с. : ил.; 24 см. - (Серия "Для профессионалов").; ISBN 978-5-4461-1371-2 : 700 экз.
3. Гольдберг Й. Нейросетевые методы в обработке естественного языка / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2019. – 282 с.: ил.
4. Kamyshev V K., Kureichik V.M., Borodyanskiy I.M., “Review of the Recommender Systems Application in Cardiology”, *Cardiometry*, 2020, no. 16, 97–105

Профессиональные базы данных, информационно-справочные системы

- 1) Единое окно доступа к образовательным ресурсам. Раздел Информатика и информационные технологии - http://window.edu.ru/catalog/p_rubr=2.2.75.6
- 2) Зональная научная библиотека УрФУ <http://lib.urfu.ru>
- 3) Научная электронная библиотека Elibrary.ru <https://www.elibrary.ru/>
- 4) Электронная библиотечная сеть "Лань" <http://e.lanbook.com/>
- 5) Портал информационно-образовательных ресурсов УрФУ <http://study.urfu.ru/>

Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

Базы данных, информационно-справочные и поисковые системы

- 1) Государственная публичная научно-техническая библиотека <http://www.gpntb.ru>
- 2) Список библиотек, доступных в Интернет и входящих в проект «Либнет» <http://www.valley.ru/-nicr/listrum.htm>
- 3) Российская национальная библиотека <http://www.rsl.ru>
- 4) Свободная энциклопедия Википедия <https://ru.wikipedia.org/>
- 5) Портал национального общества имитационного моделирования <http://simulation.su/>
- 6) Портал российской ассоциации искусственного интеллекта <https://raai.org/>

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Рекомендательные системы

Сведения об оснащении дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Подключение к сети Интернет	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES
2	Лабораторные занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Персональные компьютеры по количеству обучающихся Подключение к сети Интернет	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES
3	Самостоятельная работа студентов	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Персональные компьютеры по количеству обучающихся Подключение к сети Интернет	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES
4	Текущий контроль и промежуточная аттестация	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES

		Персональные компьютеры по количеству обучающихся Подключение к сети Интернет	
--	--	--	--

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
ПО ДИСЦИПЛИНЕ**
Рекомендательные системы

Код модуля
1164093(1)

Модуль
Рекомендательные системы

Екатеринбург

Оценочные материалы составлены автором(ами):

№ п/п	Фамилия, имя, отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Медведева Марина Александровна	кандидат физико-математических наук, доцент	Доцент	Базовая кафедра «Аналитика больших данных и методы видеоанализа»
2	Балунгу Даниель		ассистент	Базовая кафедра «Аналитика больших данных и методы видеоанализа»

Авторы:

- Медведева Марина Александровна, доцент, Базовая кафедра «Аналитика больших данных и методы видеоанализа»
- Балунгу Даниель, ассистент, Базовая кафедра «Аналитика больших данных и методы видеоанализа»

1. СТРУКТУРА И ОБЪЕМ ДИСЦИПЛИНЫ Рекомендательные системы

1.	Объем дисциплины в зачетных единицах	3	
2.	Виды аудиторных занятий	Лекции Лабораторные работы	
3.	Промежуточная аттестация	Экзамен	
4.	Текущая аттестация	Контрольная работа	1
		Домашняя работа	1

2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ (ИНДИКАТОРЫ) ПО ДИСЦИПЛИНЕ МОДУЛЯ Рекомендательные системы

Индикатор – это признак / сигнал/ маркер, который показывает, на каком уровне обучающийся должен освоить результаты обучения и их предъявление должно подтвердить факт освоения предметного содержания данной дисциплины, указанного в табл. 1.3 РПМ-РПД.

Таблица 1

Код и наименование компетенции	Индикаторы достижения компетенции	Планируемые результаты обучения	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3	4
ПК-2. Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта	ПК-2.1. Разрабатывает приложения систем искусственного интеллекта	ПК-2.1. 3-1. Знает современные языки программирования, библиотеки и программные платформы для функционального, логического, объектно-ориентированного программирования приложений систем искусственного	Лекции Лабораторные работы Домашняя работа Контрольная работа Экзамен

		интеллекта (Python, R, C++, C#) ПК-2.1. У-1. Умеет разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#)	
ПК-8. Способен создавать и внедрять одну или несколько сквозных цифровых субтехнологий искусственного интеллекта	ПК-8.1. Участвует в реализации проектов в области сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»	ПК-8.1. З-1. Знает фундаментальные правила построения рекомендательных систем и систем поддержки принятия решений, основанных на интеллектуальных принципах, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений» ПК-8.1. У-1. Умеет применять методы и подходы к планированию и реализации проектов по созданию и поддержке системы искусственного интеллекта на основе сквозной цифровой субтехнологии «Рекомендательные системы и системы поддержки принятия решений»	Лекции Лабораторные работы Домашняя работа Контрольная работа Экзамен

3. ПРОЦЕДУРЫ КОНТРОЛЯ И ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ В РАМКАХ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ МОДУЛЯ В БАЛЬНО-РЕЙТИНГОВОЙ СИСТЕМЕ (ТЕХНОЛОГИЧЕСКАЯ КАРТА БРС)

3.1. Процедуры текущей и промежуточной аттестации по дисциплине

1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0.4		
Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>домашняя работа</i>	6,14	50
<i>контрольная работа</i>	6,8	50
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 0.6		
Промежуточная аттестация по лекциям – экзамен		
Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0.4		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – не предусмотрено		
Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям– не предусмотрено		
Промежуточная аттестация по практическим/семинарским занятиям–нет		
Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям– не предусмотрено		
3. Лабораторные занятия: коэффициент значимости совокупных результатов лабораторных занятий –0.6		
Текущая аттестация на лабораторных занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>выполнение и защита лабораторных работ</i>	6,16	100
Весовой коэффициент значимости результатов текущей аттестации по лабораторным занятиям -1		
Промежуточная аттестация по лабораторным занятиям –нет		
Весовой коэффициент значимости результатов промежуточной аттестации по лабораторным занятиям – не предусмотрено		
4. Онлайн-занятия: коэффициент значимости совокупных результатов онлайн-занятий –не предусмотрено		
Текущая аттестация на онлайн-занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах

Весовой коэффициент значимости результатов текущей аттестации по онлайн-занятиям -не предусмотрено		
Промежуточная аттестация по онлайн-занятиям –нет		
Весовой коэффициент значимости результатов промежуточной аттестации по онлайн-занятиям – не предусмотрено		

3.2. Процедуры текущей и промежуточной аттестации курсовой работы/проекта

Текущая аттестация выполнения курсовой работы/проекта	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент текущей аттестации выполнения курсовой работы/проекта– не предусмотрено		
Весовой коэффициент промежуточной аттестации выполнения курсовой работы/проекта– защиты – не предусмотрено		

4. КРИТЕРИИ И УРОВНИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ МОДУЛЯ

4.1. В рамках БРС применяются утвержденные на кафедре/институте критерии (признаки) оценивания достижений студентов по дисциплине модуля (табл. 4) в рамках контрольно-оценочных мероприятий на соответствие указанным в табл.1 результатам обучения (индикаторам).

Таблица 4

Критерии оценивания учебных достижений обучающихся

Результаты обучения	Критерии оценивания учебных достижений, обучающихся на соответствие результатам обучения/индикаторам
Знания	Студент демонстрирует знания и понимание в области изучения на уровне указанных индикаторов и необходимые для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Умения	Студент может применять свои знания и понимание в контекстах, представленных в оценочных заданиях, демонстрирует освоение умений на уровне указанных индикаторов и необходимых для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Опыт /владение	Студент демонстрирует опыт в области изучения на уровне указанных индикаторов.
Другие результаты	Студент демонстрирует ответственность в освоении результатов обучения на уровне запланированных индикаторов. Студент способен выносить суждения, делать оценки и формулировать выводы в области изучения. Студент может сообщать преподавателю и коллегам своего уровня собственное понимание и умения в области изучения.

4.2 Для оценивания уровня выполнения критериев (уровня достижений обучающихся при проведении контрольно-оценочных мероприятий по дисциплине модуля) используется универсальная шкала (табл. 5).

Таблица 5

Шкала оценивания достижения результатов обучения (индикаторов) по уровням

Характеристика уровней достижения результатов обучения (индикаторов)				
№ п/п	Содержание уровня выполнения критерия оценивания результатов обучения (выполненное оценочное задание)	Шкала оценивания		
		Традиционная характеристика уровня		Качественная характеристика уровня
1.	Результаты обучения (индикаторы) достигнуты в полном объеме, замечаний нет	Отлично (80-100 баллов)	Зачтено	Высокий (В)
2.	Результаты обучения (индикаторы) в целом достигнуты, имеются замечания, которые не требуют обязательного устранения	Хорошо (60-79 баллов)		Средний (С)
3.	Результаты обучения (индикаторы) достигнуты не в полной мере, есть замечания	Удовлетворительно (40-59 баллов)		Пороговый (П)
4.	Освоение результатов обучения не соответствует индикаторам, имеются существенные ошибки и замечания, требуется доработка	Неудовлетворитель но (менее 40 баллов)	Не зачтено	Недостаточный (Н)
5.	Результат обучения не достигнут, задание не выполнено	Недостаточно свидетельств для оценивания		Нет результата

5. СОДЕРЖАНИЕ КОНТРОЛЬНО-ОЦЕНОЧНЫХ МЕРОПРИЯТИЙ ПО ДИСЦИПЛИНЕ МОДУЛЯ

5.1. Описание аудиторных контрольно-оценочных мероприятий по дисциплине модуля

5.1.1. Лекции

Изучение теоретического материала по темам/разделам лекций в соответствии с содержанием дисциплины (п. 1.2. РПД)

5.1.2. Лабораторные занятия

Примерный перечень тем

1. Постановка задачи на разработку рекомендательные системы в различных областях, таких как электронная коммерция, СМИ и социальные сети.
2. Коллаборативная фильтрация на основе сходства по пользователям (user-based) и продуктам (item-based)

3. Матричная факторизация
4. Создание гибридного рекомендателя фильмов для совместной работы с контентом с использованием глубокого обучения
5. Классификация текста с использованием Naive Bayes
6. Извлечение именованных сущностей с SpaCy
7. Оценка рекомендательных систем
8. Контекстно-зависимые рекомендации корзины и персонализированные рекомендации

LMS-платформа – не предусмотрена

5.2. Описание внеаудиторных контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля

Разноуровневое (дифференцированное) обучение.

Базовый

5.2.1. Контрольная работа

Примерный перечень тем

1. Рекомендательная система, основанная на коллаборативной фильтрации
2. Рекомендательная система, основанная на контенте
3. Рекомендательная система, основанная на знаниях

Примерные задания

1. Рекомендация элементов с использованием факторизации матрицы с функциями может (отметьте все, что применимо):
 - а) обеспечить персонализацию
 - б) захватить контекст (например, время суток)
 - в) ничего из вышеперечисленного
2. В какой из следующих ситуаций коллаборативная система фильтрации будет наиболее подходящим алгоритмом обучения (по сравнению с линейной или логистической регрессией)?
 - Вы управляете книжным интернет-магазином, и у вас есть рейтинги книг от многих пользователей. Вы хотите научиться прогнозировать ожидаемый объем продаж (количество проданных книг) в зависимости от среднего рейтинга книги.
 - Вы художник и рисуете портреты для своих клиентов. Каждый клиент получает свой портрет (себя) и дает вам обратную связь с рейтингом от 1 до 5 звезд, и каждый клиент покупает не более 1 портрета. Вы хотели бы предсказать, какую оценку даст вам ваш следующий клиент.
 - Вы управляете книжным интернет-магазином и собираете рейтинги многих пользователей. Вы хотите использовать это, чтобы определить, какие книги «похожи» друг на друга (т. е. если одному пользователю нравится определенная книга, какие другие книги ему также могут понравиться?)
 - У вас есть магазин одежды, в котором продаются джинсы многих стилей и марок. Вы собрали отзывы о различных стилях и брендах от частых покупателей, и вы хотите использовать эти обзоры, чтобы предложить этим покупателям скидки на джинсы, которые, по вашему мнению, они, скорее всего, приобретут.
 - Вы написали программное обеспечение, которое загружает новостные статьи со многих новостных сайтов. В своей системе вы также отслеживаете, какие статьи вам лично нравятся, а какие нет, и система также хранит функции этих статей (например, количество слов, имя автора). Используя эту информацию, вы хотите построить систему, чтобы попытаться найти дополнительные новые статьи, которые лично вам понравятся.

- Вы запускаете онлайн-агрегатор новостей, и для каждого пользователя вы знаете некоторое подмножество статей, которые нравятся пользователю, и какое-то другое подмножество, которое не нравится пользователю. Вы захотите использовать это, чтобы найти другие статьи, которые нравятся пользователю.

- Вы управляете книжным интернет-магазином, и у вас есть рейтинги книг от многих пользователей. Для каждого пользователя вы хотите порекомендовать другие книги, которые ему понравятся, основываясь на его собственных оценках и оценках других пользователей.

3. Вы управляете киноимперией и хотите создать систему рекомендаций фильмов, основанную на коллаборативной фильтрации. Было три популярных обзорных веб-сайта (которые мы назовем А, В и С), на которые пользователи переходили, чтобы оценить фильмы, и вы только что приобрели все три компании, которые управляют этими веб-сайтами. Вы хотели бы объединить наборы данных трех компаний, чтобы создать единую / унифицированную систему. На веб-сайте А пользователи оценивают фильм как имеющий от 1 до 5 звезд. На веб-сайте В пользователи ранжируются по шкале от 1 до 10, допускаются десятичные значения (например, 7,5). На сайте С рейтинги от 1 до 100. У вас также достаточно информации, чтобы идентифицировать пользователей/фильмы на одном веб-сайте с пользователями/фильмами на другом веб-сайте. Какое из следующих утверждений верно?

- Вы можете объединить три набора данных в один, но сначала вы должны нормализовать оценки каждого набора данных (скажем, изменить масштаб оценок каждого набора данных до диапазона 0-1).

- Вы можете объединить все три обучающих набора в один, если вы выполняете нормализацию среднего значения и масштабирование функций после объединения данных.

- Предполагая, что в одной базе данных есть хотя бы один фильм/пользователь, который также не отображается во второй базе данных, нет надежного способа объединить наборы данных из-за отсутствия данных.

- Невозможно объединить данные этих веб-сайтов. Необходимо создать три отдельные рекомендательные системы.

- Вы можете объединить три набора данных в один, но сначала вы должны нормализовать каждый набор данных отдельно, вычтя среднее значение, а затем разделив на $(\max - \min)$, где \max и \min (5-1) или (10-1) или (100-1) для трех веб-сайтов соответственно.

4. Что такое рекомендательная система, основанная на содержании?

- Рекомендательная система, основанная на контенте, пытается рекомендовать товары пользователям на основе их профиля, основанного на их предпочтениях и вкусах.

- Рекомендательная система, основанная на содержании, пытается рекомендовать элементы на основе сходства между пунктами

- Все вышеперечисленное.

5. Что такое рекомендательная система на основе памяти?

- В подходе, основанном на памяти, рекомендательная система создается с использованием методов машинного обучения, таких как регрессия, кластеризация, классификация и т. д.

- В подходе, основанном на памяти, разрабатывается модель пользователей в попытке узнать их предпочтения

- В подходе, основанном на памяти, мы используем весь набор данных пользовательских элементов для создания системы рекомендаций.

6. Что означает «холодный старт» в коллаборативной фильтрации?

- Сложность рекомендаций, когда у нас недостаточно оценок в наборе данных пользовательского элемента

- Сложность в рекомендации, когда у нас есть новый пользователь, и мы не можем сделать для него профиль, или когда у нас есть новый элемент, который еще не получил никакого рейтинга.

- Сложность рекомендаций, когда количество пользователей или элементов увеличивается, а объем данных расходуется, поэтому алгоритмы начнут страдать от падения производительности.

LMS-платформа – не предусмотрена

5.2.2. Домашняя работа

Примерный перечень тем

1. Разработка рекомендательной системы на основе коллаборативной фильтрации.
2. Проектирование алгоритма под рекомендательную систему.
3. Оценка рекомендательных систем.

Примерные задания

1. Разработать рекомендательную систему на основе коллаборативной фильтрации. Выполнить задание с использованием Python, Pandas, SKLearn, Matplotlib, Seaborn, Jupyter Notebook.

2. Источником кода с комментариями : <http://blog.ethanrosenthal.com/2016/01/09/explicit-matrix-factorization-sgd-als/>

- Сравнить качество рекомендаций и время выполнения ALS (без базовых предикторов - without biases) и SGD (без базовых предикторов) для различных параметров регуляризации (например, 0,001, 0,01,...) и числа факторов (например, 1, 10, 25, 50,...). Для SGD необходимо усовершенствовать код.

- Исправить код ALS так, чтобы была возможность производить расчеты с базовыми предикторами (biases). Неплохая инструкция тут: <http://activisiongamescience.github.io/2016/01/11/Implicit-Recommender-Systems-Biased-Matrix-Factorization/>.

- Сравнить все 4 подхода по мере MSE, времени на обучение (training phase) и тестирование (inference phase). Результаты, помимо шагов выполнения в ipython, необходимо представить в сводной таблице.

Источник данных: MovieLens 100k. Может быть выбран альтернативный, но не меньший по числу оценок, либо пользователей и предметов (items).

3. Сравнение методов коллаборативной фильтрации по сходству пользователей и по сходству объектов

1. Требуется реализовать вычисление ошибки [MAE](http://www.recsyswiki.com/wiki/Mean_absolute_error) и [RMSE](http://www.recsyswiki.com/wiki/Root_mean_square_error-mean-square_deviation) на тестовых данных [Movie Lens](<http://grouplens.org/datasets/movielens/>).

В качестве данных обучения можно использовать файлы с расширением base, а тестирование качества провести на файле test: пары файлов u1.base и u1.test, ..., u5.base и u5.test. Каждая пара — это разбиение 80%/20% данных для всех пользователей u на обучающие и тестовые данные.

2. Для каждого метода (user-based и item-based) постройте графики зависимости [MAE](http://www.recsyswiki.com/wiki/Mean_absolute_error) и [RMSE](http://www.recsyswiki.com/wiki/Root_mean_square_error-mean-square_deviation) от числа соседей (диапазон от 1 до 100 с разумным шагом).

3. Если качество предсказаний слишком низкое ($MAE > 2,0$), то попробуйте формулы 2.6 и 2.7 из обзора <http://files.grouplens.org/papers/FnT%20CF%20Recsys%20Survey.pdf>.

Можно использовать альтернативные формулы для исходной модели $r_{u,i} = k \sum \limits_{u' \in U} \operatorname{sim}(u, u') r_{u', i}$ (случай user-based модели):

$$r_{u,i} = \frac{1}{|N|} \sum \limits_{u' \in U} r_{u', i}$$

$$r_{u,i} = \bar{r}_u + \sum_{u' \in U} \text{sim}(u, u') (r_{u',i} - \bar{r}_{u'})$$

где $k = 1 / \sum_{u' \in U} \text{sim}(u, u')$

4. Сравните подходы на основе полученных результатов по аналогии с пунктами 1 и 2.
5. Как изменяется величина MAE (RMSE) от числа выдаваемых рекомендаций (top-n): $n \in \{1, 3, 5, 10, 15, 20, 30, 40, 50, 100\}$?
6. Как Вы считаете, какие фильмы чаще рекомендуются -- популярные с высокими оценками или редкие (те, которые редко оцениваются) с высокими оценками?
7. Что делать, если соседей (то есть похожих на целевого пользователя или конкретный товар) мало? Нужно/можно ли как-то учитывать достоверность таких рекомендаций?

LMS-платформа – не предусмотрена

5.3. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине модуля

5.3.1. Экзамен

Список примерных вопросов

1. Типы рекомендательных систем
 - 1) Фильтрация на основе контента
 - 2) Совместная фильтрация
2. Рекомендательные системы на основе контента
 - 1) Анализ документов с помощью TF-IDF
 - 2) Создание векторизатора TF-IDF
 - 3) Вычисление косинусного подобия – скалярного произведения нормализованных векторов
 - 4) Преимущества рекомендательных систем, основанных на контенте
 - 5) Ограничения рекомендательных систем, основанных на контенте
3. Рекомендательные системы совместной фильтрации
 - 1) Совместная фильтрация на основе памяти
 - 2) Рекомендательная система совместной фильтрации на основе моделей
4. NLP и рекомендательные системы

LMS-платформа – не предусмотрена

5.4 Содержание контрольно-оценочных мероприятий по направлениям воспитательной деятельности

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения	Контрольно-оценочные мероприятия
Профессиональное воспитание	профориентационная деятельность	Технология формирования уверенности и готовности к самостоятельной успешной	ПК-2	ПК-2.1. У-1.	Лекции Экзамен Домашняя работа

		профессиональн ой деятельности Технология самостоятельно й работы			
--	--	---	--	--	--