

Кафедра инженерной кибернетики

Учебная дисциплина
«Экспертные и рекомендательные системы»

ЛАБОРАТОРНАЯ РАБОТА №3

*«Разработка и реализация методов коллаборативной фильтрации
(Collaborative Filtering, CF) в асинхронном чат-боте
(на примере мессенджера Telegram)»*

бакалавриат по направлению *01.03.04 прикладная математика*
(семестр VII; 2025÷2026 уч. год)

Содержание

Введение.....	3
1. Задание на лабораторную работу	4
1.1. Основная цель работы.....	4
1.2. Общие сведения о коллаборативной фильтрации.....	4
1.3. Общие сведения о чат-ботах.....	6
2. Выполнение лабораторной работы.....	9
3. Защита лабораторной работы	11
4. Требования к содержанию и оформлению отчета	12
Приложение. Титульный лист отчета по лабораторной работе	15

Введение

Настоящий документ содержит задание для **лабораторной работы №3** (далее – л/р №3) по дисциплине «**Экспертные и рекомендательные системы**» для учащихся бакалавриата по направлению подготовки 01.03.04 *прикладная математика* кафедры инженерной кибернетики НИТУ «МИСИС».

Основным направлением работ учащихся при выполнении л/р №3 являются процессы совместного использования технологий разработки асинхронных чат-ботов и методов *коллаборативной фильтрации*, используемых при создании *рекомендательных систем*.

Режим выполнения учащимися л/р №3 (индивидуальный, командный) определяет преподаватель. Выбор средств разработки должен быть согласован с преподавателем, ведущим лабораторные работы по учебной дисциплине.

Внимание! При разработке программного кода заданного вида коллаборативной фильтрации учащимся нельзя использовать готовые сторонние программные средства (программные библиотеки, фреймворки и т.п.), т.е. учащийся должен осуществить программирование полностью самостоятельно.

Каждый учащийся при выполнении лабораторной работы обязан соблюдать интеллектуальные, авторские и смежные права и лицензионные соглашения соответствующих правообладателей при использовании в процессах разработки своего программного обеспечения любых выбранных технических средств и инструментов, а также научного инструментария.

1. Задание на лабораторную работу

1.1. Основная цель работы

Основная цель лабораторной работы – выработать у учащихся:

- понимание принципов и особенностей использования такого подхода к созданию рекомендательных алгоритмов, как метод коллаборативной фильтрации,
- понимание принципов и особенностей использования в области рекомендательных технологий, таких числовых мер (статистик), как косинусное сходство и коэффициент корреляции Пирсона,
- начальные навыки владения процессом создания асинхронных чат-ботов для мессенджера *Telegram*©.

1.2. Общие сведения о коллаборативной фильтрации

Товарные рекомендации – один из самых популярных методов продвижения товаров в розничной торговле. Они упрощают посетителям поиск новых, интересных именно им товаров. На сегодняшний день эффективные рекомендации (как товарные, так и контентные) являются ключевым фактором роста конверсий и дохода коммерческих предприятий.

Онлайн-рекомендации – это продукт сложных алгоритмических решений: система анализирует различные взаимодействия пользователя и находит возможности для его вовлечения (*engagement*) в реальном времени. При выборе правильной рекомендательной стратегии важно учесть множество факторов. Одни стратегии (подходы) делают больший упор на товары и формируют рекомендации на базе общих рейтингов (товары в тренде, популярные статьи и т.п.) или в контексте конкретных страниц (статьи, похожие текущую; товары, которые часто покупают с текущим, и т.п.). Однако есть и подходы (стратегии), которые задействуют всю мощь персонализации и генерируют рекомендации на основании индивидуальных предпочтений и сходств (*affinities*) пользователя.

Коллаборативная фильтрация (англ. *collaborative filtering*) – один из методов построения прогнозов (рекомендаций) в рекомендательных системах, использующий известные предпочтения (оценки) группы пользователей для прогнозирования неизвестных предпочтений другого пользователя. Суть такого подхода заключается в том, что рекомендательная система собирает данные о пользователях, выявляет подобия (*similarities*) между ними и, основываясь на этой информации, делает более точные товарные рекомендации.

При использовании коллаборативной фильтрации анализируется только реакция пользователей на объекты: оценки, которые выставляют пользователи объектам. Оценки могут быть как **явными** (пользователь выставляет оценки объектам), так и **неявными** (например, количество просмотров одного ролика). Чем больше оценок собирается, тем точнее получатся рекомендации. Таким образом, пользователи помогают друг другу в фильтрации объектов, поэтому такой метод называется также **совместной** (по-английски: **коллаборативной** – *collaborative*) **фильтрацией**.

Необходимо отметить, что коллаборативная фильтрация вырабатывает рекомендации, основанные на модели предшествующего поведения пользователя. Эта модель может быть построена исключительно на основе поведения данного пользователя или – что более эффективно – с учетом поведения других пользователей со сходными характеристиками. Коллаборативная фильтрация использует два разных типа входных данных: множество пользователей и множество объектов интереса. Отношения между пользователями и объектами интереса обычно выражаются при помощи оценок, предоставляемых пользователями, и использующихся в последующих сессиях для прогнозирования оценок, которые пользователь мог бы поставить неоцененным объектам интереса.

Выделяют несколько подходов применения коллаборативной фильтрации, среди которых основными являются:

- а) ***user-based*** collaborative filtering – это подход построения множества «*рекомендаций, основанных на пользователях*», используя который находятся пользователи со схожими предпочтениями и одному из них то рекомендуется то, что понравилось другим похожим пользователям;

б) *item-based* collaborative filtering – это подход построения множества «*рекомендаций, основанных на объектах интереса (товарах, продуктах)*» где анализируется схожесть между объектами интереса на основе пользовательских оценок и, если многие пользователи одинаково оценили два товара, то эти товары считаются похожими. Иными словами, рекомендации строятся на основе схожести между товарами или контентом. Если конкретный пользователь предпочитал определенный товар, то ему будут рекомендованы похожие товары, которые другие пользователи с похожими интересами также предпочли.

1.3. Общие сведения о чат-ботах

Чат-бот (англ. *chatbot*; также - виртуальный собеседник, программа-собеседник) – это программа для мобильного устройства и/или настольного ПК, которая выясняет потребности пользователей, а затем помогает удовлетворить их. В чат-ботах реализуется имитация общения с пользователем с «живым» сотрудником некоторой организации. Такое общение ведется с помощью текста или голоса и используются такие программы как альтернатива переписке с живым оператором или звонку сотруднику организации.

Чат-боты упрощают взаимодействие между пользователем некоторого интернет-сервиса или мессенджера (например, Telegram) и организацией и помогают экономить средства. Быстрый поиск ответа на вопрос, смена персональных данных, устранение мелких неполадок в приложении, оформление заявки на покупку продукта – с этим чат-бот вполне может справиться самостоятельно. При необходимости он перенаправляет запрос вместе с полученными данными на нужного специалиста.

Чат-бот — это компьютерная программа, которая относится к категории интеллектуальных систем, и представляет собой электронного (виртуального) «собеседника», созданного для решения типовых задач: поиск информации, ответы на вопросы клиентов-посетителей различных интернет-сервисов, маркетплейсов т.п. электронных площадок, выполнение простых процедур и действий в самых разных сферах, например:

- банковская сфера – консультация, техническая поддержка, а также выполнение стандартных процедур (например, открытие банковского вклада) и простых финансовых транзакций;
- операторы связи – техническая поддержка;
- страхование – помощь в заполнении форм и заявок, консультация по страхованию;
- онлайн-торговля (маркетплейсы) – консультирование по доставке, оплате, адресам точек выдачи;
- здравоохранение – предоставление медицинских материалов, первичная консультация и сбор анамнеза;
- туризм – агрегатор предложений, рассылка горящих туров, бронирование;
- образование – рассылки, приглашение на вебинары;
- государственные услуги – сбор жалоб, быстрый доступ к публичным данным;
- HR-службы – подбор подходящих резюме, автоматизация задач внутри компании.

Задачи, которые могут решать чат-боты, также могут быть самыми разными:

- распознавание и интерпретация запросов,
- консультация по вопросам банковского обслуживания и совершение финансовых транзакций,
- бронирование столиков в ресторанах,
- запись на посещение в разных организациях (медицинские услуги, салоны красоты, фитнес-центры и т.п.);
- решение стандартных задач техподдержки компьютерного и сетевого оборудования (например, при взаимодействии с провайдером))
- и многое другое,

Существуют разнообразные виды чат-ботов, которые можно классифицировать: по платформе внедрения, технологии разработки, способам организации процессов общения с пользователями (синхронные и асинхронные) и по различным аспектам функциональности.

Основное отличие синхронного чат-бота от асинхронного заключается в том, что в **синхронных** – каждый запрос от пользователя блокирует выполнение всего чат-бота до момента выдачи ответа на запрос, а в **асинхронных** такой блокировки нет и чат-бот может обрабатывать несколько запросов параллельно.

Асинхронность является предпочтительным режимом работы, так как повышает производительность и улучшает реактивность бота, чтобы он мог эффективнее справляться с различными задачами. Однако реализации асинхронности требует более высокой квалификации от разработчика и больших затрат при создании системы.

В л/р №3 учащийся должен освоить технологию создания синхронных чат-ботов с использованием с использованием специализированного инструментария – **конструктора**.

2. Выполнение лабораторной работы

Основная задача. Учащемуся необходимо разработать заданный вариант реализации метода коллаборативной фильтрации (Collaborative Filtering – CF) в асинхронном чат-боте для мессенджера Telegram, функционирующего на условно-реальных исходных данных. При выполнении запрещается использовать готовые методы для коллаборативной фильтрации и меры сходства из каких-либо библиотек.

Варианты реализации:

1) User-Based Collaborative Filtering с использованием в качестве меры сходства – *косинусного сходства*¹.

2) Item-Based Collaborative Filtering с использованием в качестве меры сходства – *коэффициент корреляции Пирсона*².

3) Item-Based Collaborative Filtering с использованием в качестве меры сходства – *косинусного сходства*.

4) User-Based Collaborative Filtering с использованием в качестве меры сходства – *коэффициент корреляции Пирсона*;

Программный код должен быть: модульным, иметь комментарии и аннотацию типов, явно типизирован.

Разрешается использовать как реальные данные для осуществления рекомендации, например датасет: <https://www.kaggle.com/datasets/prajitdatta/movielens-100k-dataset/data> , так и сгенерированные пользователем.

Основные этапы выполнения лабораторной работы учащимся следующие:

1) Определение преподавателем конкретного варианта реализации л/р №3 (см. выше).

2) Выбор средств для создания чат-бота.

¹ Косинусное сходство – это мера сходства между двумя векторами предгильбертового пространства, которая используется для измерения косинуса угла между ними

² Коэффициент корреляции Пирсона (или линейный коэффициент корреляции) – это статистика, которая измеряет величину линейной связи (корреляцию) между двумя переменными. Принимает значения от -1 до +1. Значение коэффициента +1 означает наличие полной положительной линейной связи, а значение -1 – наличие полной отрицательной линейной связи

- 3) Проектирование программной архитектуры CF-модели и чат-бота: назначение, основная и дополнительная функциональность.
- 4) Непосредственная разработка программного кода для заданного метода коллаборативной фильтрации.
- 5) Непосредственная разработка программного кода асинхронного чат-бота.
- 6) Тестирование (проверка работоспособности и корректности работы) созданного программного кода.
- 7) Выполнение вычислений по заданным исходным данным.
- 8) Анализ полученных результатов.
- 9) Подготовка доклада по результатам работы (до 5 мин.).
- 10) Защита результатов работы преподавателю.

Работа выполняется индивидуально.

Внимание! Допускается выполнение л/р №3 в команде по согласованию с преподавателем.

Дополнительная информация.

Для начального знакомства с особенностями асинхронного программирования на языке Python рекомендуется ознакомиться с информацией по ссылке: <https://school.kontur.ru/publications/2567>

P.S. Язык программирования можно использовать любой.

3. Защита лабораторной работы

Выполненная учащимся (группой учащихся) лабораторная работа проходит процедуру защиты. При проведении защиты лабораторной работы учащийся (группа учащихся):

- демонстрирует работоспособность созданного программного обеспечения и безошибочность реализации заявленной функциональности,
- рассказывает про:
 - решаемую задачу,
 - выбранные и реализованные технологии в программном коде,
 - проблемы, которые возникли в ходе выполнения работы,
 - использованные средства программирования для создания чат-бота и методов коллаборативной фильтрации (назначения, возможности, ограничения и проч.),
- представляет подготовленный отчет по **л/р №3**.

В процессе и (или) после демонстрации преподаватель задает учащемуся (членам команды) вопросы по различным аспектам выполнения **л/р №3** и полученным в ходе её выполнения результатам. В случае наличия существенных замечаний и (или) выявления ошибок в работе созданного программного обеспечения и тексте отчета защита **л/р №3** прекращается до полного исправления выявленных ошибок и устранения сделанных замечаний.

После того как преподавателем будет окончательно принято созданное в **л/р №3** программное обеспечение (чат-бот), результаты его работы и отчет, учащийся обязан сдать¹ преподавателю оформленный отчет по **л/р №3**.

Отчет должен содержать: титульный лист, историческая справка, описание решаемой задачи с помощью конструктора, скриншоты работы конструктора, демонстрация вложенности, преимущества и недостатки данного конструктора, вывод. Форма и режим сдачи отчетных материалов объявляются преподавателем отдельно.

¹ По решению преподавателя отчетные материалы могут либо сдаваться в распечатанном виде (на твердом носителе) либо загружаются учащимся в электронном виде в LMS Moodle

4. Требования к содержанию и оформлению отчета

1) Общие требования.

Язык отчета – русский.

Текст отчета должен быть проверен на наличие и не должен содержать орфографических и синтаксических ошибок.

2) Требования к содержанию отчета.

Отчет по лабораторной работе оформляется в соответствии со следующими требованиями к содержанию.

- **Титульный лист** (оформляется в соответствии с **приложением**).
- **Введение**

Описываются: основная цель и задачи лабораторной работы.

- **Краткое (не более 2 стр.) описание решаемой задачи по интеллектуальной обработке и/или анализу информации**

Описываются: основное содержание решаемой задачи, что является исходными данными (с примерами), что является результатом решения задачи и дополнительные сведения, раскрывающие особенности как самой задачи, так и процесса её решения (при необходимости).

- **Выбранные средства для разработки программного обеспечения**

Описываются: используемые ИИ-сервис/систем и конкретный ИИ-функционал (модули; подсистемы; библиотеки; платформы; API и т.п.), а также использованные для создания программного обеспечения языки программирования, сторонние библиотеки и фреймворки, среды разработки (IDE), СУБД (если используется) и т.п.

- **Описание и примеры исходных данных**

Учащийся должен указать фразы текста на естественном языке, которые были использованы в качестве стартовых фраз для модели.

- **Результаты**

Иллюстрированное описание процесса решения задачи с использованием созданного программного обеспечения. Описание выполняется в форме последовательности скриншотов экранных форм с сопроводительным краткими текстовыми комментариями, иллюстрирующих основной процесс работы созданного приложения при решении задачи для заданного числа примеров исходных данных.

– **Анализ полученных результатов работы**

В этой части отчета учащийся (команда учащихся) должен привести развернутый обоснованный анализ, содержащий наиболее значимые выводы и оценку полученных результатов л/р №3.

- **Выводы по лабораторной работе.**
- **Список использованных источников**
- **Приложения (при необходимости)**

2) Требования к оформлению отчета.

Текст отчета, таблицы, графики и диаграммы, формулы, другие специальные обозначения должны быть отредактированы, а также отформатированы единым образом. Сокращения слов, за исключением общепринятых, не допускаются.

Параметры страницы.

- Размер листа – формат «А4»;
- Поля: левое: 25 мм, правое: 10 мм, верх и низ: 15 мм;
- Печать текста на листе – односторонняя.

Форматирование текста.

- Шрифты:
 - ♦ основной текст – Times New Roman Cyr, 12 пт.;
 - ♦ названия глав (разделов) отчета – Times New Roman Cyr, 14 пт., набор прописными буквами с полужирным выделением (bold) и выравниванием по центру страницы;
 - ♦ заголовки пунктов и подпунктов – Times New Roman Cyr, 14 пт. с полужирным выделением (bold);

- ♦ подписи рисунков – Times New Roman Cyr, 11 пт.

–Абзацы:

- ♦ первая строка – отступ: 1,25 см;
- ♦ межстрочный интервал – 1,5;
- ♦ выравнивание абзаца – по ширине;
- ♦ интервал после абзаца обычного текста – 6 пт.;
- ♦ интервал до и после заголовков пунктов и подпунктов – 12 пт.;

–Нумерация страниц: внизу по центру; титульный лист является первой страницей отчета и не нумеруется.

–Нумерация рисунков и таблиц – сквозная целочисленная. Если таблица и/или рисунок в отчете единственные, то тогда они не нумеруются.

–Все таблицы и рисунки должны иметь название.

–Все графики и графические изображения, отражающие какие-либо функциональные зависимости, должны содержать подписи по всем обозначенным осям (абсцисс, ординат, аппликат), включающие в себя: название; единицу измерения соответствующей величины; числовые отметки на оси, характеризующие масштаб отображаемой величины.

–Если рисунок или таблица размещается на странице (страницах), для которой задана «альбомная» ориентация листа, то их размещение должно быть таким, чтобы правильное расположение объекта относительно читателя достигалось поворотом страницы из нормального положения документа на 90° по часовой стрелке.

–Размещаемые в отчете рисунки не должны вызывать затруднений для рассмотрения, а текстовая информация на рисунках должна быть читабельной.

Кафедра инженерной кибернетики

ОТЧЕТ

ПО

ЛАБОРАТОРНОЙ РАБОТЕ №3

*«Разработка и реализация методов коллаборативной фильтрации
(Collaborative Filtering, CF) в асинхронном чат-боте
(на примере мессенджера Telegram)»*

учебная дисциплина

«Экспертные и рекомендательные системы»

Студент: ФИО и группа

Преподаватель: Хонер П.Д.

Оценка: _____

Дата защиты: _____

2025 г.