THE UNIVERSITY OF DANANG

DANANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

GRADUATION PROJECT THESIS

MAJOR: INFORMATION TECHNOLOGY

SPECIALTY: SOFTWARE ENGINEERING

PROJECT TITLE:

BOOK MANGEMENT APPLICATION INTEGRATES DEEP LEARNING

Instructor: **Ph.D. PHAM CONG THANG**

Student: **NGUYEN VAN MINH NHAT**

Student ID: **102200223**

Class**:** **20TCLC_DT5**

Da Nang, 6/2024

DA NANG UNIVERSITY

UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

SOCIALIST REPUBLIC OF VIETNAM

Independence - Freedom - Happiness

GRADUATION PROJECT COMMENT

1. **General information:**

1. Student name: Nguyen Van Minh Nhat
2. Class: 20TCLC_DT5 Student ID: 102200223
3. Topic title: Book management application integrates deep learning
4. Instructor: Ph.D. Pham Cong Thang Academic title/ degree: Ph.D.

2. **Reviews of graduation project**

- About the urgency, novelty, usability of the topic: (2 points)

………………………………………………………………………………………

………………………………………………………………………………………

- About the results of solving the tasks required by the project: (4 points)

………………………………………………………………………………………

………………………………………………………………………………………

- About the form, structure and layout of the graduation project: (2 points)

………………………………………………………………………………………

………………………………………………………………………………………

- The topic includes scientific value/article/problem solving of the enterprise or school: (1 point)

………………………………………………………………………………………

………………………………………………………………………………………

- Existing shortcoming need to be supplemented or modified:

………………………………………………………………………………………

………………………………………………………………………………………

3. **Spirit and attitude of the student (1 point):**

………………………………………………………………………………………

………………………………………………………………………………………

4. **Evaluation:**

- Evaluation point: …./10
- Suggest: Defense permitted/ Edit to defend/ Defense not permitted

Da Nang, date……month……. 2024

Instructor

Ph.D. Pham Cong Thang

SUMMARY

Topic title: Book management system integrated deep learning

Student name: Nguyễn Văn Minh Nhật

Student ID: 102200223 Class: 20TCLC_DT5

The Book Management System is a web application designed to assist bookstores in managing their book sales online. This application offers a range of features including managing book prices, authors, languages, and more. Beyond these fundamental capabilities, the application integrates advanced deep learning technologies to enhance user experience and boost bookstore revenue.

One of the key features is **semantic search**. Unlike traditional keyword-based search methods, semantic search allows users to find books using natural language sentences. Even if the search query does not match the book title exactly, the system can identify and retrieve books that are contextually relevant to the search sentence.

Another significant feature is the **recommender system**. This tool assists users who are uncertain about their next purchase by suggesting books they might be interested in. By analyzing user preferences and behaviors, the recommender system increases the likelihood of users finding their next favorite book, thereby improving sales and user satisfaction.

This project aims to create an efficient, user-friendly platform for bookstore managers and their customers.

PREFACE

I would like to express my sincere and deep gratitude to the lecturers in the Faculty of Information Technology, as well as all the lecturers at the University of Science and Technology – University of Da Nang, who have guided, taught, and imparted their valuable knowledge and experience to me throughout my studies and research at the university.

I would like to express my heartfelt thanks and appreciation to Associate Professor Dr. Pham Cong Thang, who has guided and helped me step by step during the process of completing my graduation thesis. Thanks to his guidance, I was able to complete the thesis on schedule and accumulate valuable knowledge.

I would like to send my greatest thanks to my father, mother, and family. My parents and family have always been by my side, providing tireless motivation and strong spiritual support, helping me overcome difficulties to complete this thesis.

Although I have tried my best to complete the thesis, due to limited time and knowledge, there will inevitably be certain shortcomings. I earnestly hope for the sympathy, guidance, and suggestions from the esteemed lecturers and all my friends to make my work more complete.

Once again, I sincerely thank you all

ASSURRANCE

I hereby declare that:

1. The contents of this thesis were carried out by me under the direct guidance of Professor Pham Cong Thang.
2. All references used in the thesis are clearly cited with the author's name, the title of the work, and the time and place of publication.
3. If there are any improper copies or violations of training regulations, I take full responsibility.

Student Performed

TABLE OF CONTENT

LIST OF TABLES

PICTURES

LIST OF SYMBOL, ACRONYM

Symbol	Acronym	Explaination
API	Application Programing Interface	A set of rules and protocols that allows different software applications to communicate with each other. APIs define the methods and data formats that applications can use to request and exchange information.
Restful	Representational State Transfer	A type of web API that adheres to the principles of REST, an architectural style for designing networked applications
AI	ArtiArtificial Intelligence	Refers to the simulation of human intelligence in machines that are programmed to think and learn. These machines can perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation
BERT	Bidirectional Encoder Representations from Transformers	A groundbreaking natural language processing (NLP) model developed by Google.
SBERT	Sentence Bidirectional Encoder Representations from Transformers	An extension of the BERT model specifically designed to generate semantically

		meaningful sentence embeddings
NLP	Natural language processing	A field of artificial intelligence that focuses on the interaction between computers and humans through natural language. The goal of NLP is to enable computers to understand, interpret, and generate human language in a way that is both meaningful and useful

INTRODUCTION

1. Purpose

- Help customer can find their book with only smartphone or computer by find it in our website

- Help admin users manage everything about their website such as: books, authors, language, category, reviews, …

2. Target

- Building a website for customer with simple and attractive user interface, and help user can buy their favorite book

- Building semantic search for user can search books they want and integrate it into system

- Building recommender system using deep learning, training model, evaluate and add it into system

- Building admin site so that admin can manage their books, and their customer so that they can control everything they want

3. Scope and objects of research

The topic target to 2 objects:

- Peoples who want to find and buy their book in internet with smartphone or computer
- Peoples who want a website to manage their book and sale it to customer

Chapter 1: OVERVIEW

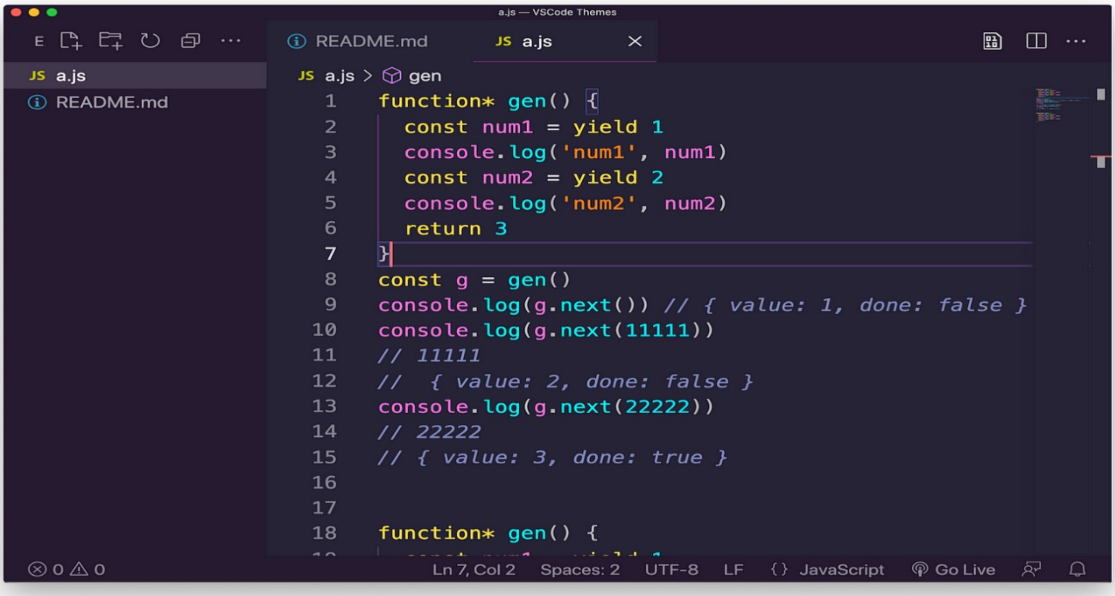
1.1. Introduce topic

Nowadays, the rapid development of the internet has significantly changed user behavior, especially in terms of buying and selling. People no longer need to spend time going out to find stores that sell the items they want. Instead, they can go online to search for and purchase items, making shopping more convenient. This applies to buying books as well. Now, everyone can quickly and conveniently find the books they want on websites. Additionally, this provides advantages that traditional bookstores cannot, such as book recommendations based on the behavior of many users or finding books based on search queries, even if the query does not match the book's title exactly.

1.2. Tools

1.2.1. Visual studio code

Visual Studio Code, also commonly referred to as VS Code, is a source-code editor developed by Microsoft for Windows, Linux, macOS and web browsers. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded version control with Git



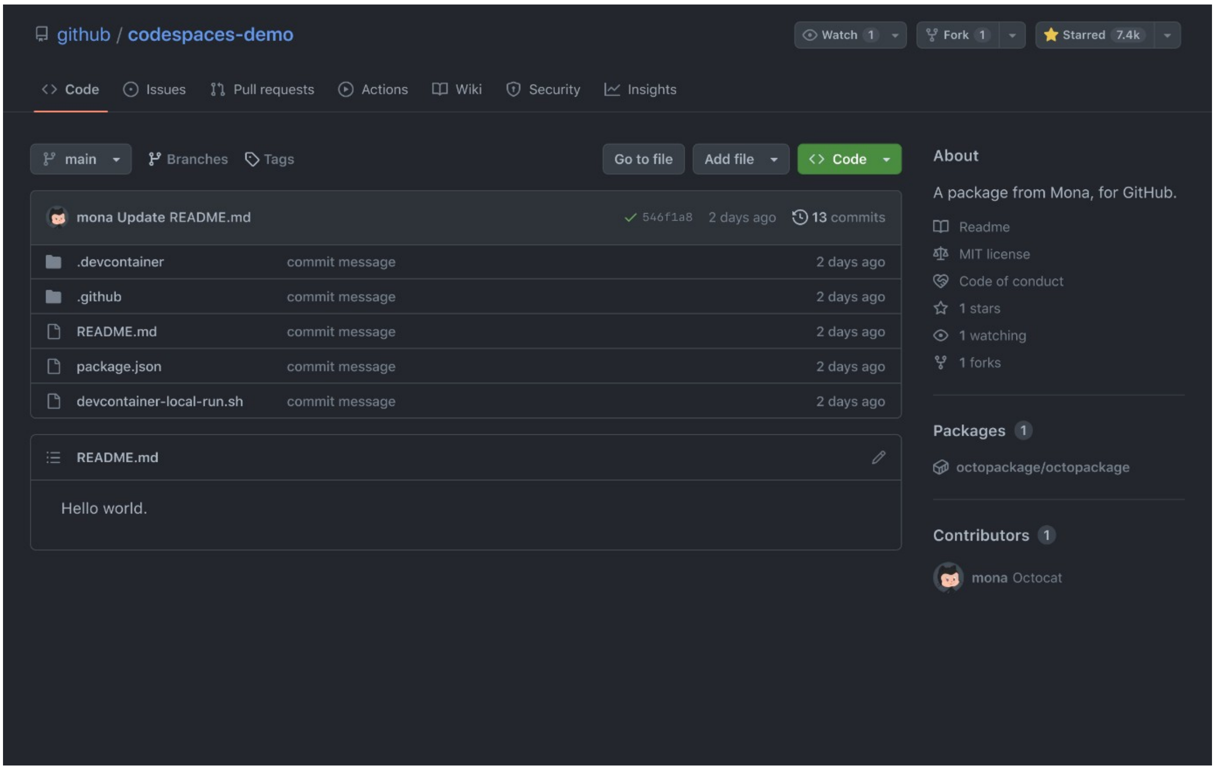
```
function* gen() {
  const num1 = yield 1
  console.log('num1', num1)
  const num2 = yield 2
  console.log('num2', num2)
  return 3
}
const g = gen()
console.log(g.next()) // { value: 1, done: false }
console.log(g.next(11111))
// 11111
//  { value: 2, done: false }
console.log(g.next(22222))
// 22222
// { value: 3, done: true }


function* gen() {
```

Picture 1. 1 Visual studio code GUI 1

1.2.3. Git/GitHub

- **Git:** is a distributed version control system that tracks versions of files. It is often used to control source code by programmers collaboratively developing software.
- **GitHub:** is a developer platform that allows developers to create, store, manage and share their code. It uses Git software, providing the distributed version control of Git plus access control, bug tracking, software feature requests, task management, continuous integration, and wikis for every project



mona Update README.md ✓ 546f1a8 2 days ago 13 commits

.devcontainer	commit message	2 days ago
.github	commit message	2 days ago
README.md	commit message	2 days ago
package.json	commit message	2 days ago
devcontainer-local-run.sh	commit message	2 days ago

README.md

Hello world.

Picture 1. 2 GitHub GUI

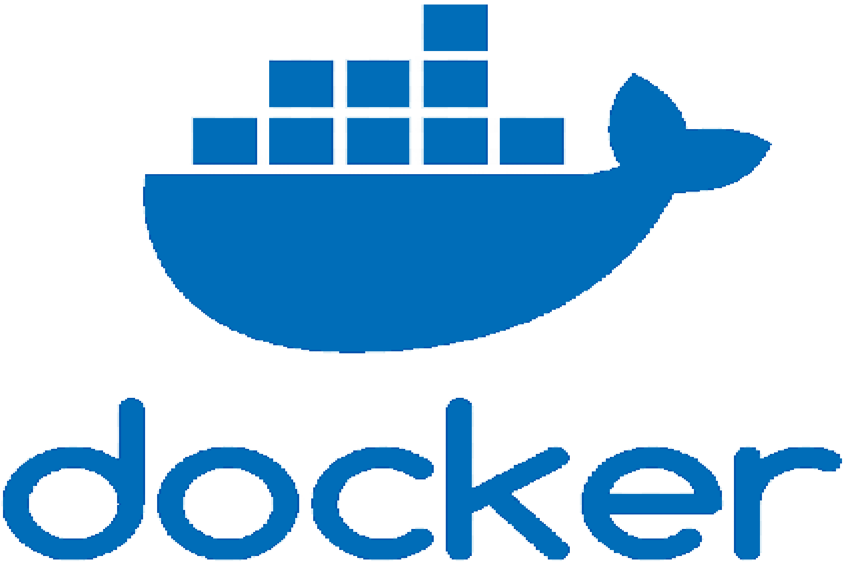
1.2.4. PostgreSQL

PostgreSQL, also known as Postgres, is a free and open-source relational database management system emphasizing extensibility and SQL compliance

Picture 1. 3 Introduce PostgreSQL

1.2.5. Docker

Docker is a set of platforms as a service product that use OS-level virtualization to deliver software in packages called containers. The service has both free and premium tiers. The software that hosts the containers is called Docker Engine. It was first released in 2013 and is developed by Docker, Inc.



Picture 1. 4 Introduce docker

1.2.6. Google Colab

Colab is a hosted Jupyter Notebook service that requires no setup to use and provides free access to computing resources, including GPUs and TPUs. Colab is especially well suited to machine learning, data science, and education.

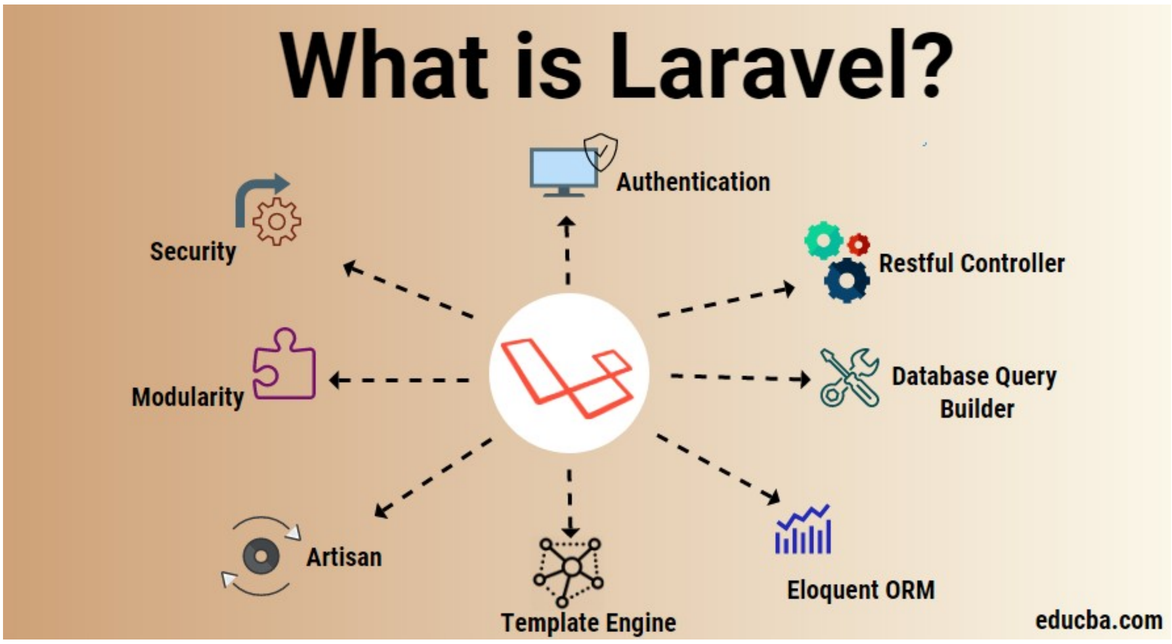
Picture 1. 5 Introduce Google Colab

Chapter 2: THEORETICAL BASIS

2.1. PHP/Laravel

- **PHP** is a general-purpose scripting language geared towards web development. It is used most in web development. Beside that it is simple and easy to learn and understand, so that everyone can use to create their own website in the short of time

- **Laravel** a powerful and elegant PHP framework, has garnered widespread acclaim in the web development community for its robust features, intuitive syntax, and developer-friendly approach, enabling the creation of modern, scalable web applications with ease through its comprehensive toolset, including built-in support for features such as routing, authentication, ORM (Object-Relational Mapping) with Eloquent, MVC (Model-View-Controller) architecture, Blade templating engine for dynamic content rendering, artisan command-line interface for efficient scaffolding and management tasks, seamless integration with popular front-end frameworks like Vue.js and React, robust security implementations, including CSRF (Cross-Site Request Forgery) protection and encryption, extensive testing capabilities with PHPUnit, and a vibrant ecosystem supported by an active community and numerous libraries, making it an ideal choice for developers aiming to build sophisticated, maintainable applications swiftly and effectively.



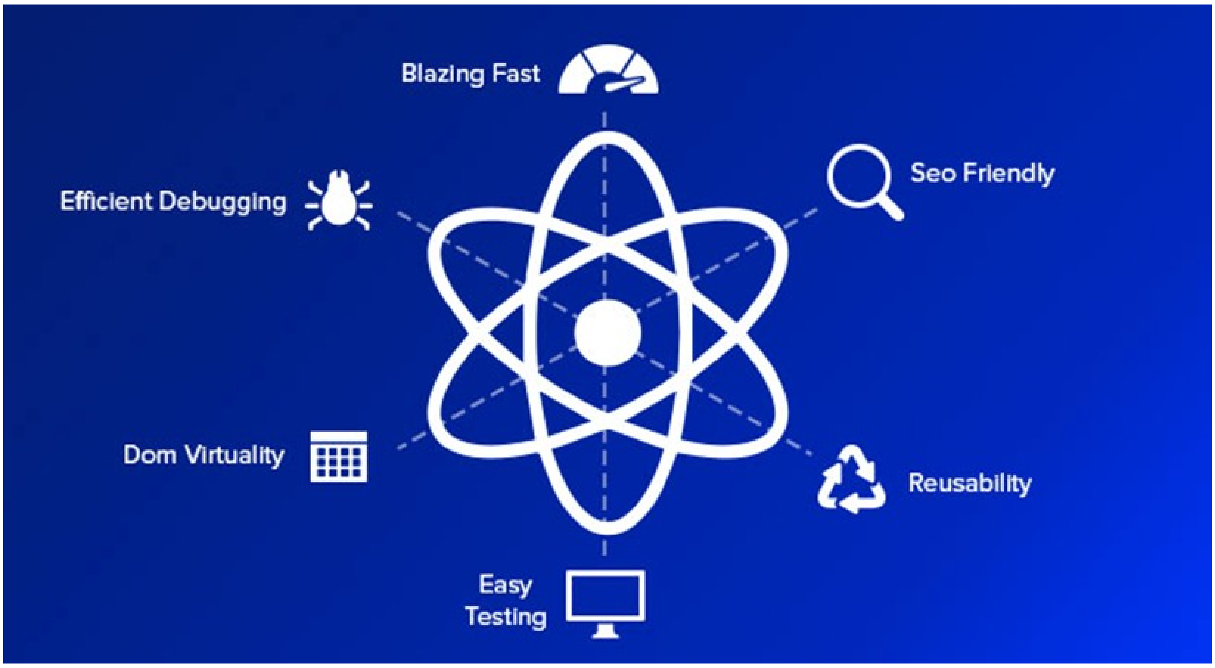
Picture 2. 1 Laravel feature

2.2. ReactJS

React (React.js or ReactJS) is a JavaScript library, open-source and directly developed by Facebook, designed to efficiently build user interfaces. With ReactJS, one can consider using it to design websites or utilize React Native for easy development of mobile applications. The core objective of React.js is to create fast web applications with minimal code, capable of operating smoothly, reliably, and easily scalable when maintenance is required. With React, instead of focusing on the entire website, you can divide it into smaller parts, making tasks simpler, more convenient, and timesaving. ReactJS is widely used on major popular websites such as WhatsApp, Airbnb, Instagram, Facebook, Netflix, and many others.

A prominent feature of ReactJS is its focus on solving the view layer, thus creating components as the main building blocks of applications. Each functionality in ReactJS is divided into different components for efficient use, facilitating the reuse of components with similar functions and avoiding potential code duplication.

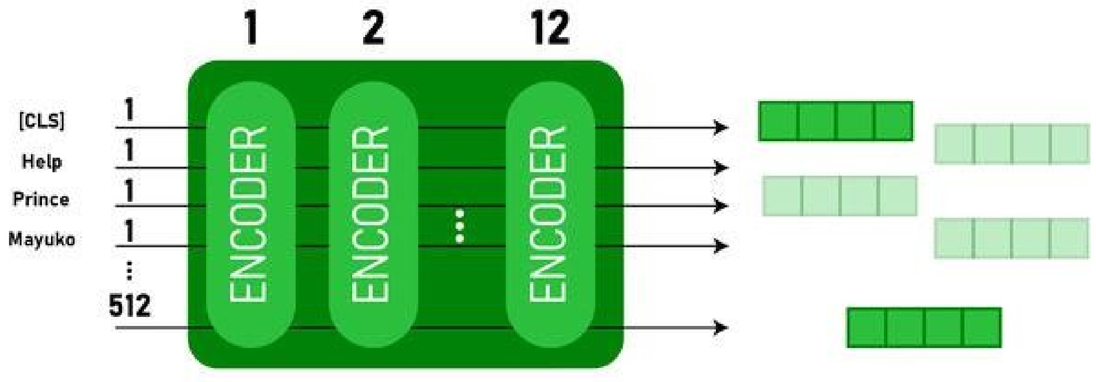
ReactJS employs Virtual DOM (Document Object Model), which enables easy manipulation with the Real DOM. Consequently, changes in the actual DOM are swiftly reflected in the Virtual DOM, enhancing application performance, ensuring smooth operations, and achieving desired results efficiently.



Picture 2. 2 ReactJs advamtages

2.3. BERT

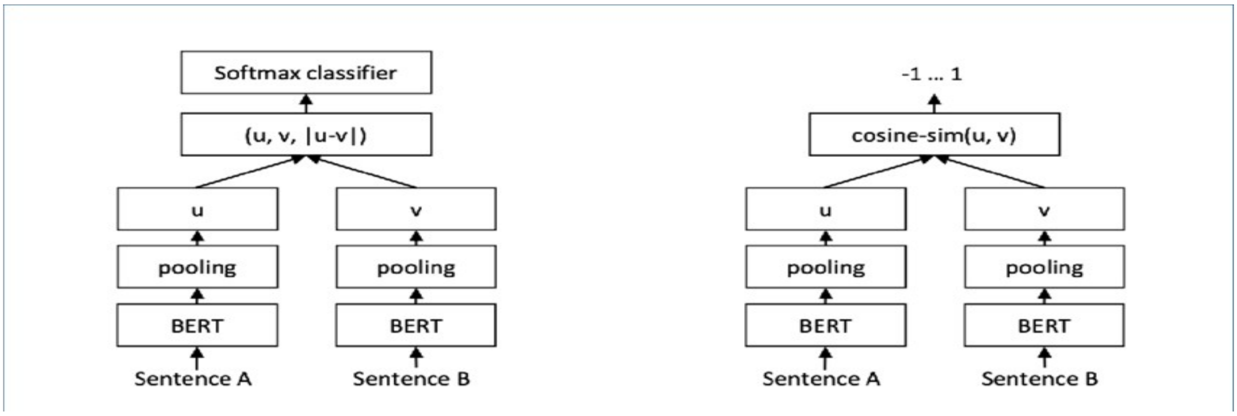
BERT, or Bidirectional Encoder Representations from Transformers, is a state-of-the-art natural language processing model developed by Google AI Language that revolutionized the field by leveraging the bidirectional training of Transformers, enabling deep contextual understanding of language through pre-training on vast amounts of text data, subsequently fine-tuned for specific NLP tasks, thereby achieving significant advancements in tasks such as question answering, sentiment analysis, and language translation, due to its ability to capture complex linguistic patterns and relationships, making it highly versatile and effective across a wide range of applications in both research and industry.



Picture 2. 3 BERT architecture

2.4. SBERT

SBERT, or Sentence-BERT, represents a specialized variant of BERT tailored specifically for sentence embeddings. Unlike traditional BERT, which focuses on token-level representations, SBERT aims to generate embeddings that capture the semantic meaning of entire sentences. This is achieved through training strategies that optimize for similarity between semantically similar sentences in embedding space, enabling tasks like semantic textual similarity, clustering, and retrieval to benefit from more meaningful representations. SBERT's innovations lie in its ability to efficiently encode sentences into fixed-size vectors that preserve semantic relationships, thus enhancing the performance of various downstream NLP applications that require understanding of sentence-level semantics.



Picture 2. 4 SBERT architecture

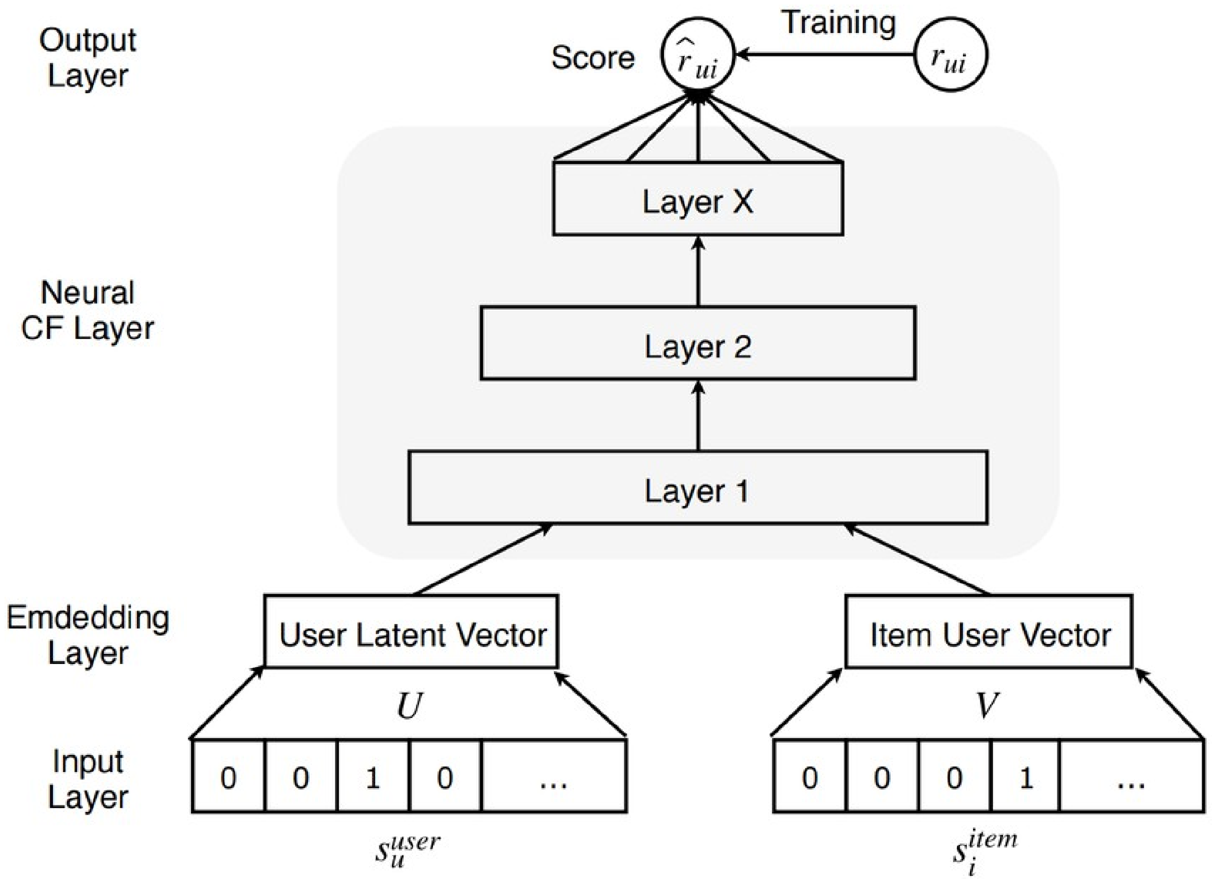
2.6. Neural collaborative filtering

Neural collaborative filtering (Neural CF) is an advanced approach to collaborative filtering that leverages neural networks to learn the underlying patterns in user-item interactions for recommendation systems. Unlike traditional collaborative filtering methods that rely on matrix factorization or similarity metrics, Neural CF uses deep learning models to capture intricate relationships between users and items directly from raw data.

In Neural CF, the model typically consists of an embedding layer for users and items, which maps them into a latent space where interactions between users and items can be modeled. These embeddings are learned jointly with the neural network during training, allowing the model to capture nuanced user preferences and item characteristics.

One popular architecture for Neural CF is the use of multi-layer perceptrons (MLPs) or deep neural networks (DNNs). These networks take user and item embeddings as input and pass them through multiple layers of non-linear transformations to predict the likelihood of a user interacting with or preferring an item. The model is trained using techniques like stochastic gradient descent (SGD) or its variants to minimize the prediction error between the model's outputs and observed interactions in the training data.

Neural CF offers several advantages over traditional collaborative filtering methods, including the ability to handle sparse and high-dimensional data more effectively, capture complex patterns in user-item interactions, and integrate additional features such as user demographics or contextual information. These advancements often lead to more accurate and personalized recommendations, enhancing user satisfaction and engagement in various applications such as e-commerce, entertainment platforms, and social networks.



Picture 2. 5 Neural collaborative filtering architecture

Chapter 3: SYSTEM ANALYSIS AND DESIGN

3.1. Functional analysis

3.1.1 Functions for admin

- **Login function**

  This function allows admin user log in to system to operate features and can be log out any time
- **Add Book function**

  This function helps admin users add books into websites with various fields like book cover, book title, ...
- **Add Language function**

  This function help admin users add languages, so that when creating books, admin users can choose which languages of those books
- **Add Author function**

  This function help admin users add authors, so that when creating books, admin users can choose which author of those books

3.1.2 Functions for user

- **Register function**

  This function allows users register their account, so that they can use that account to log in into system
- **Login function**

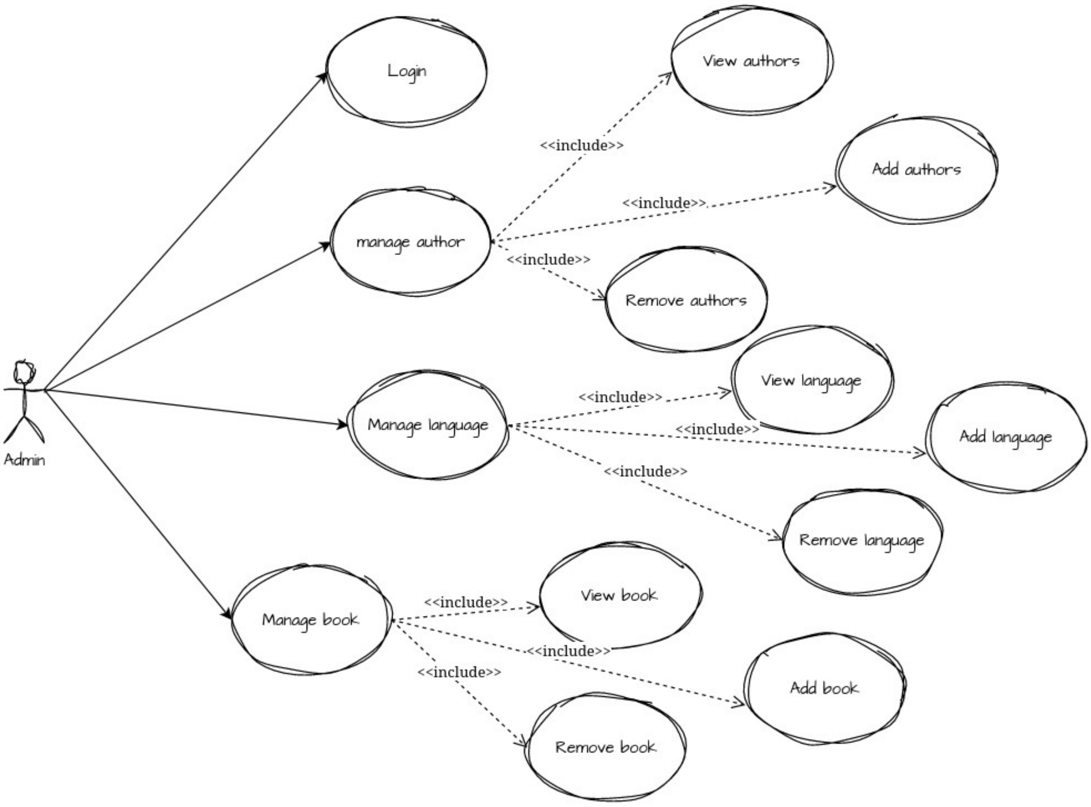
  This function allows user log in to website so that they can have many feature the guest not has
- **Search function**

  This function provides 2 modes for users to find books they want; the first one is keyword search which will find title base on keywords. Second one is semantic search which will find book has similarity to search query
- **Purchase book**

  This function allow user can purchase books which they have already added to cart

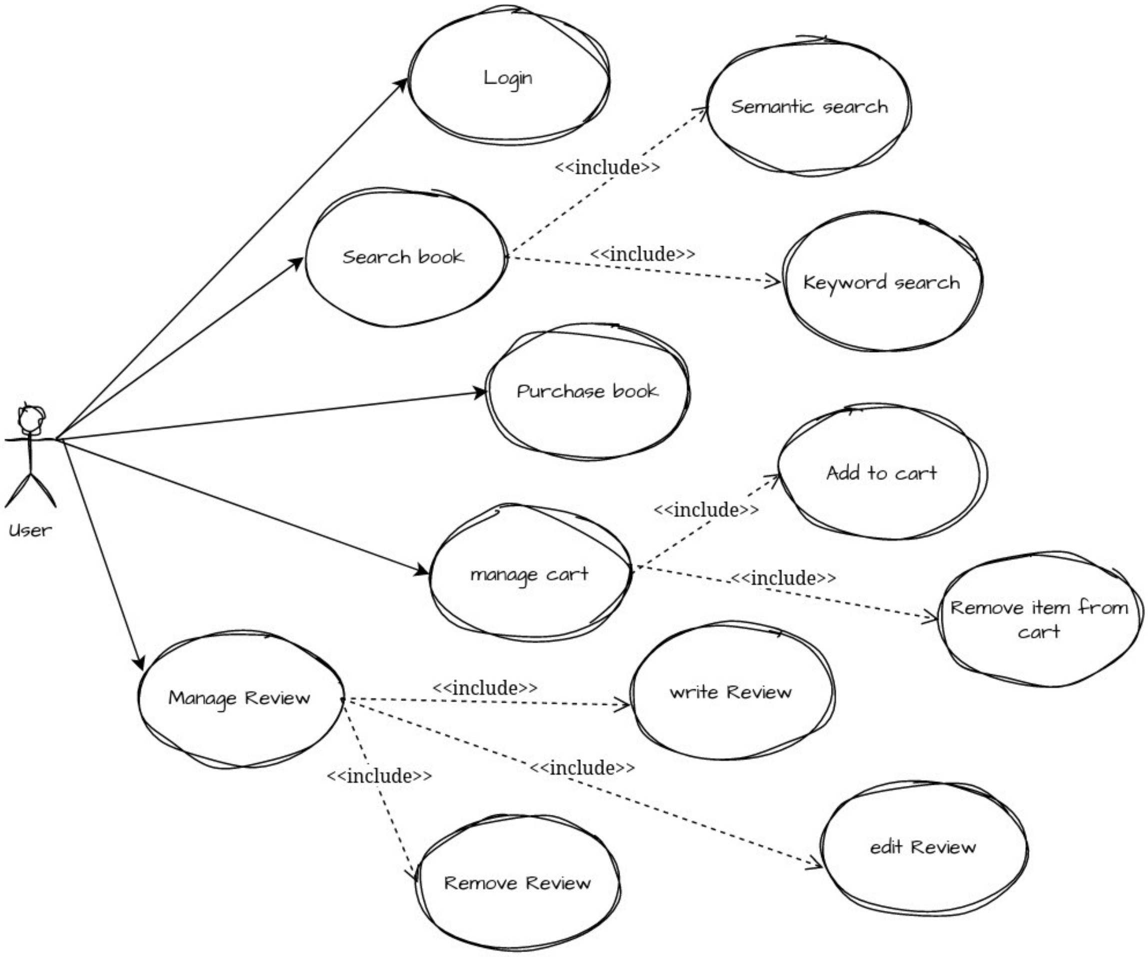
3.2. Use case diagram

3.2.1 Use case for admin



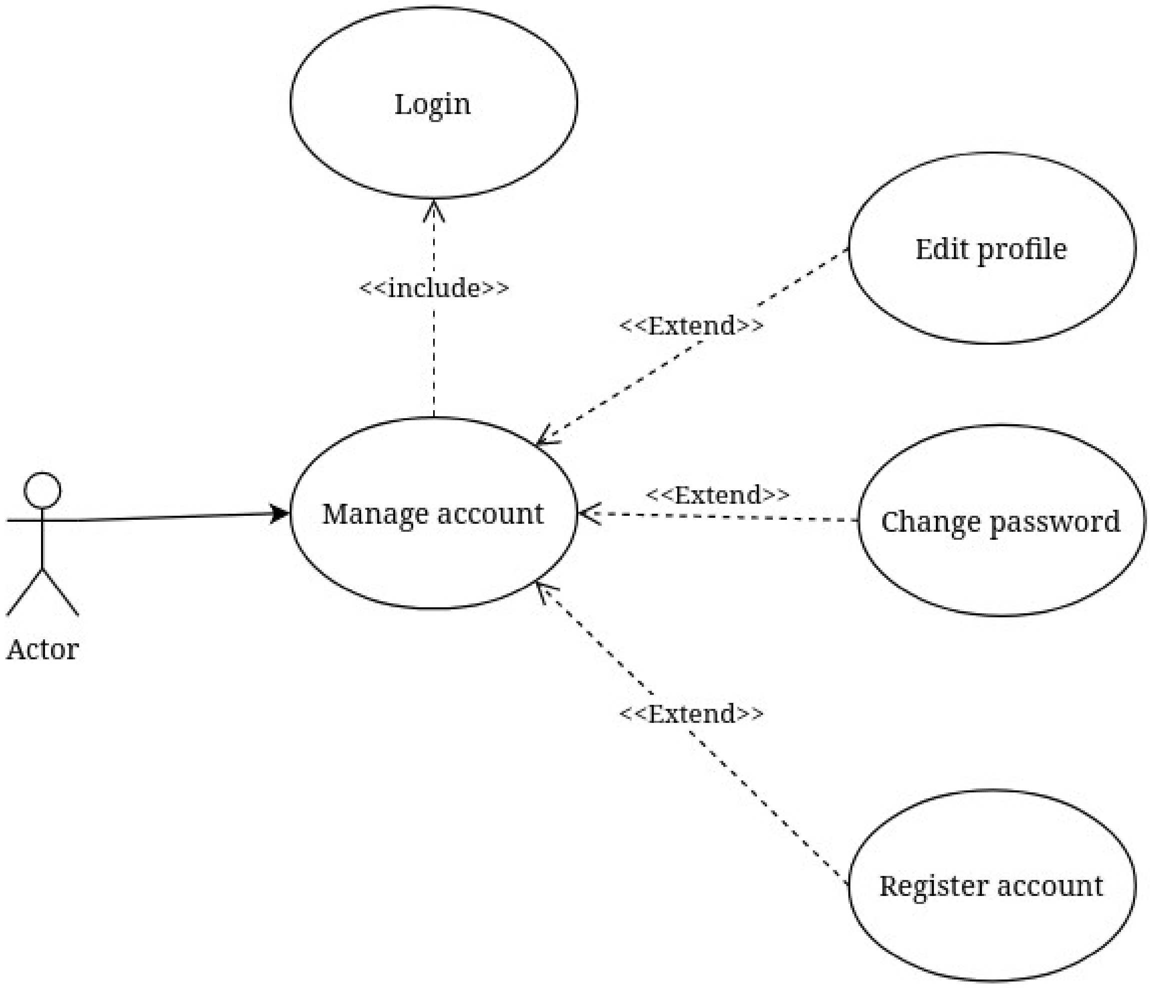
Picture 3. 1 Admin use case diagram

3.2.2 Use case for customer



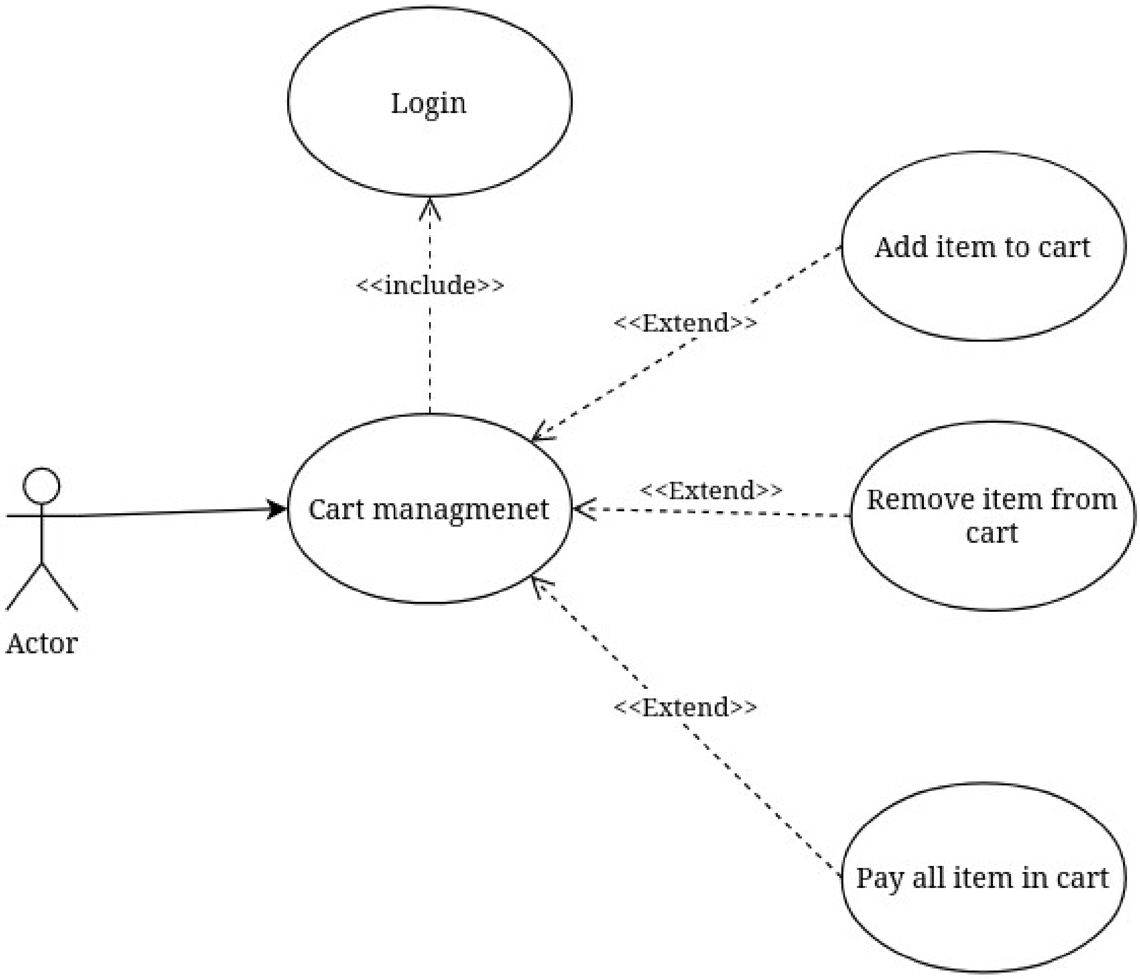
Picture 3. 2 Customer use case diagram

3.2.3 Use case for account management



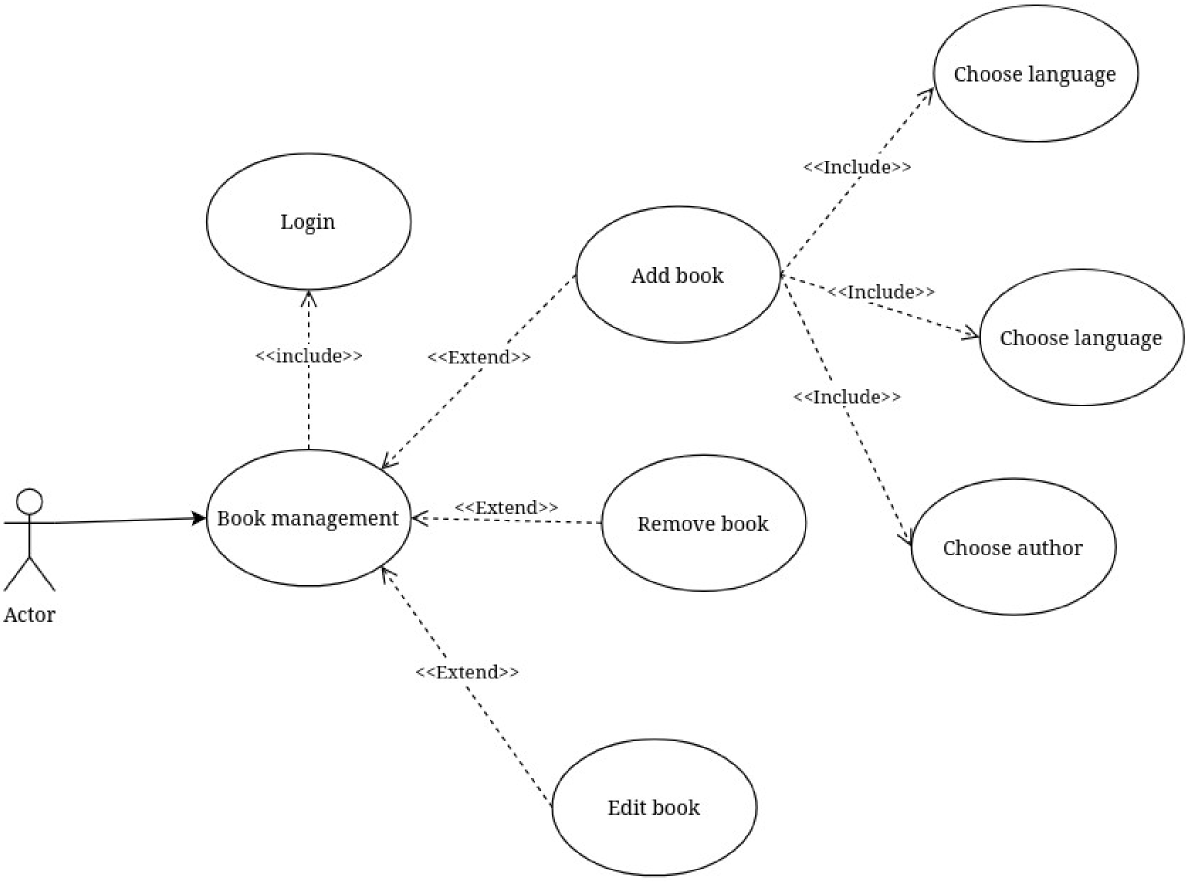
Picture 3. 3 Account management use case diagram

3.2.4 Use case for cart management



Picture 3. 4 Cart management use case diagram

3.2.5 Use case for book management



Picture 3. 5 Cart management use case diagram

3.3. Use case specification

Use case name	Register
Actor	User
Description	This function allows user to create new account and use this account to log in to system
preconditions	User has not logged into system
Path	1. Click Register button 2. Enter correct Register 3. information 4. Click submit

Table 3. 1 Register function specification

3.3.2. Login function

Use case name	Login
Actor	User/Admin
Description	This function allows user login into system
preconditions	User has not logged into system
Path	1. Click login button 2. Enter correct login information 3. Click submit button

Table 3. 2 Login function specification

3.3.3 Purchase book function

Use case name	Purchase book
Actor	User
Description	This function allows user choose book and purchase it
preconditions	User has already logged into system
Path	1. Choose Book want to buy 2. Add to cart or click purchase directly 3. Click Purchase 4. Choose payment method 5. Pay the order

Table 3. 3 Purchase book function specification

3.3.4 Search book function

Use case name	Search book
Actor	User
Description	This function allows user find books they want
preconditions	
Path	5. Click search bar 6. Choose search mode 7. Enter search query 8. Press Enter

Table 3. 4 Search book function specification

3.3.5 Add book function

Use case name	Add book
Actor	Admin
Description	This function allows admin to add new books into site so that customer can view and purchase it
preconditions	Admin user has already logged into system
Path	1. Choose book management section on sidebar 2. Click add button 3. Enter all required fields and information 4. Click submit button

Table 3. 5 Add book function specification

3.3.6 Add language function

Use case name	Add language
Actor	Admin
Description	This function allows admin to add new language so that this language will appear when user create new or edit book
preconditions	Admin user has already logged into system
Path	1. Choose language management section on sidebar 2. Click add language 3. Enter all required fields and information 4. Click submit button

Table 3. 6 Add language function specification

3.3.7 Add author function

Use case name	Add language
Actor	Admin
Description	This function allows admin to create new authors so that authors can appear when admin create or edit books
preconditions	Admin user has already logged into system
Path	1. Choose author management section on sidebar 2. Click add author 3. Enter all required fields and information 4. Click submit button

Table 3. 7 Add author function specification

3.4 Database design

3.4.1 Database table

- **user:** Table includes all user information such as name, password, email address, phone number, ...

- **category: Table include all categories of books**

- **author:** Table includes all author information such as name, avatar, ...

- **language:** Table includes all languages information such as name, icon, ...

- **Books:** Table includes all books information such as: title, cover, price, quantity, description, ...

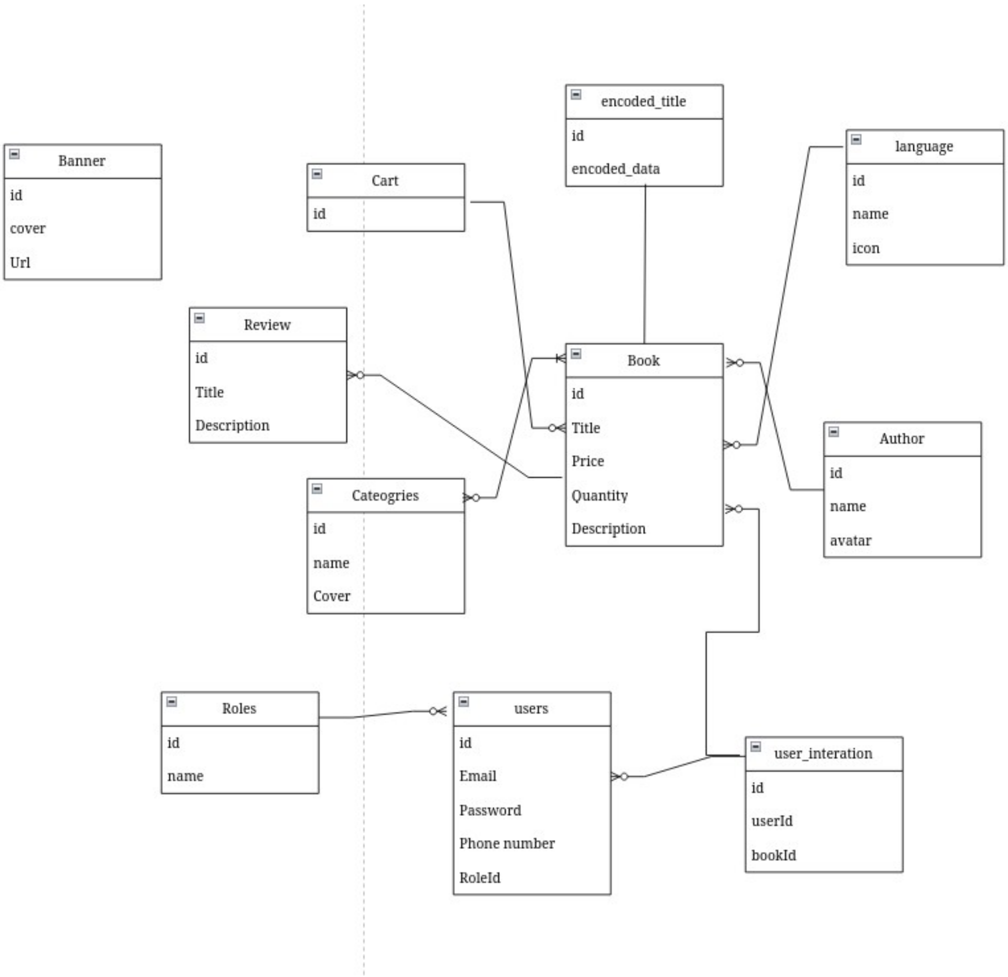
- **Carts:** Table includes all items which user add into

- **Reviews:** Table include all review for each book which user write

- **encoded_title:** Table included all encoded title using for semantic search

- **user_interation: T**able included all item that user has clicked on it, this using for recommendation system

3.4.2 Database table detail

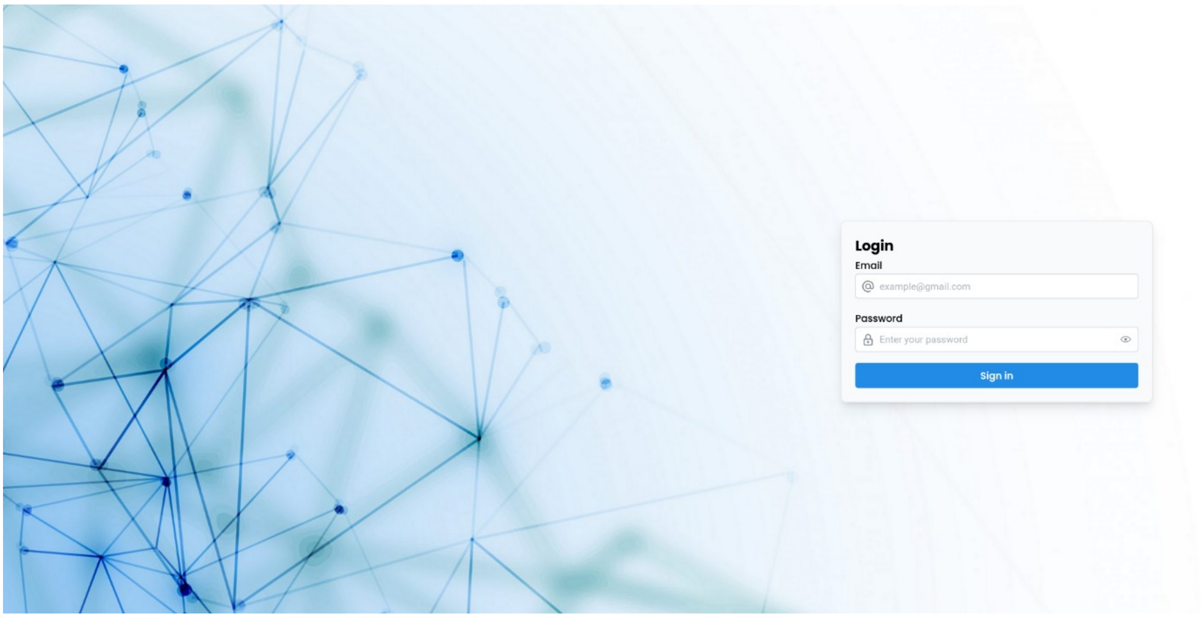


Picture 3. 6 Database design diagram

Chapter 4: IMPLEMENTATION AND EVALUATION

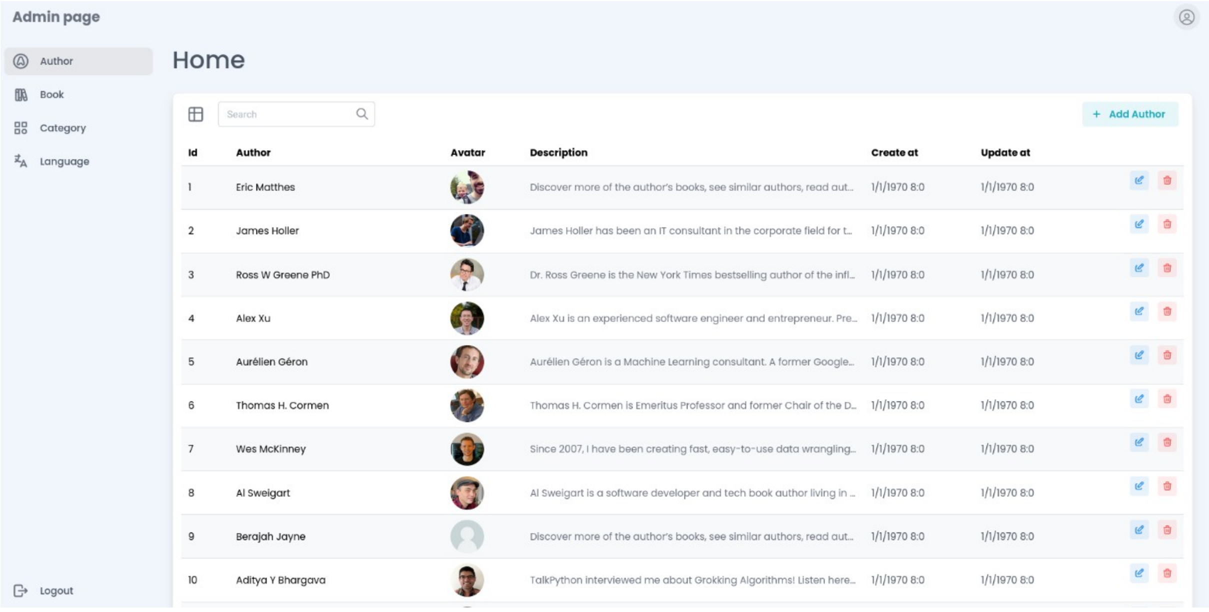
4.1 Admin page

4.1.1 Login page



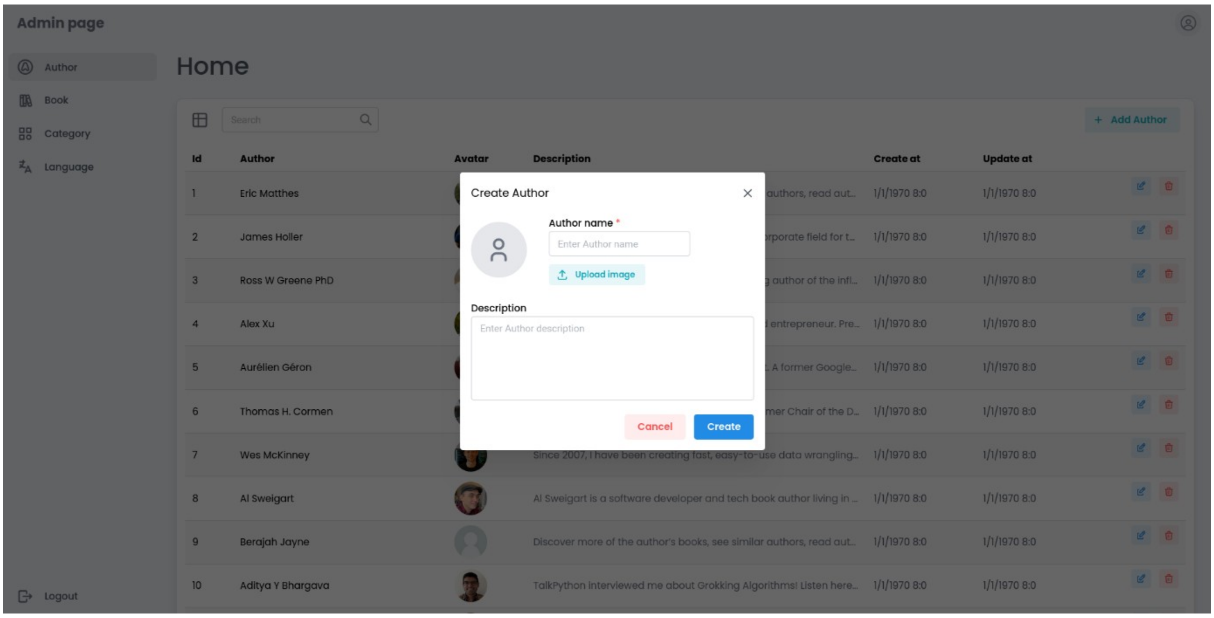
Picture 4. 1 Login page

4.1.2 Author management page



Id	Author	Avatar	Description	Create at	Update at
1	Eric Matthes		Discover more of the author's books, see similar authors, read aut...	1/1/1970 8:0	1/1/1970 8:0
2	James Holler		James Holler has been an IT consultant in the corporate field for t...	1/1/1970 8:0	1/1/1970 8:0
3	Ross W Greene PhD		Dr. Ross Greene is the New York Times bestselling author of the infl...	1/1/1970 8:0	1/1/1970 8:0
4	Alex Xu		Alex Xu is an experienced software engineer and entrepreneur. Pre...	1/1/1970 8:0	1/1/1970 8:0
5	Aurélien Géron		Aurélien Géron is a Machine Learning consultant. A former Google...	1/1/1970 8:0	1/1/1970 8:0
6	Thomas H. Cormen		Thomas H. Cormen is Emeritus Professor and former Chair of the D...	1/1/1970 8:0	1/1/1970 8:0
7	Wes McKinney		Since 2007, I have been creating fast, easy-to-use data wrangling...	1/1/1970 8:0	1/1/1970 8:0
8	Al Sweigart		Al Sweigart is a software developer and tech book author living in ...	1/1/1970 8:0	1/1/1970 8:0
9	Berajah Jayne		Discover more of the author's books, see similar authors, read aut...	1/1/1970 8:0	1/1/1970 8:0
10	Aditya Y Bhargava		TalkPython interviewed me about Grokking Algorithms! Listen here...	1/1/1970 8:0	1/1/1970 8:0

Picture 4. 2 Author management page



Id	Author	Avatar	Description	Create at	Update at
1	Eric Matthes		authors, read aut...	1/1/1970 8:0	1/1/1970 8:0
2	James Holler		rporate field for t...	1/1/1970 8:0	1/1/1970 8:0
3	Ross W Greene PhD		author of the infl...	1/1/1970 8:0	1/1/1970 8:0
4	Alex Xu		entrepreneur. Pre...	1/1/1970 8:0	1/1/1970 8:0
5	Aurélien Géron		. A former Google...	1/1/1970 8:0	1/1/1970 8:0
6	Thomas H. Cormen		mer Chair of the D...	1/1/1970 8:0	1/1/1970 8:0
7	Wes McKinney		Since 2007, I have been creating fast, easy-to-use data wrangling...	1/1/1970 8:0	1/1/1970 8:0
8	Al Sweigart		Al Sweigart is a software developer and tech book author living in ...	1/1/1970 8:0	1/1/1970 8:0
9	Berajah Jayne		Discover more of the author's books, see similar authors, read aut...	1/1/1970 8:0	1/1/1970 8:0
10	Aditya Y Bhargava		TalkPython interviewed me about Grokking Algorithms! Listen here...	1/1/1970 8:0	1/1/1970 8:0

Create Author ×

Author name *

Enter Author name

Upload image

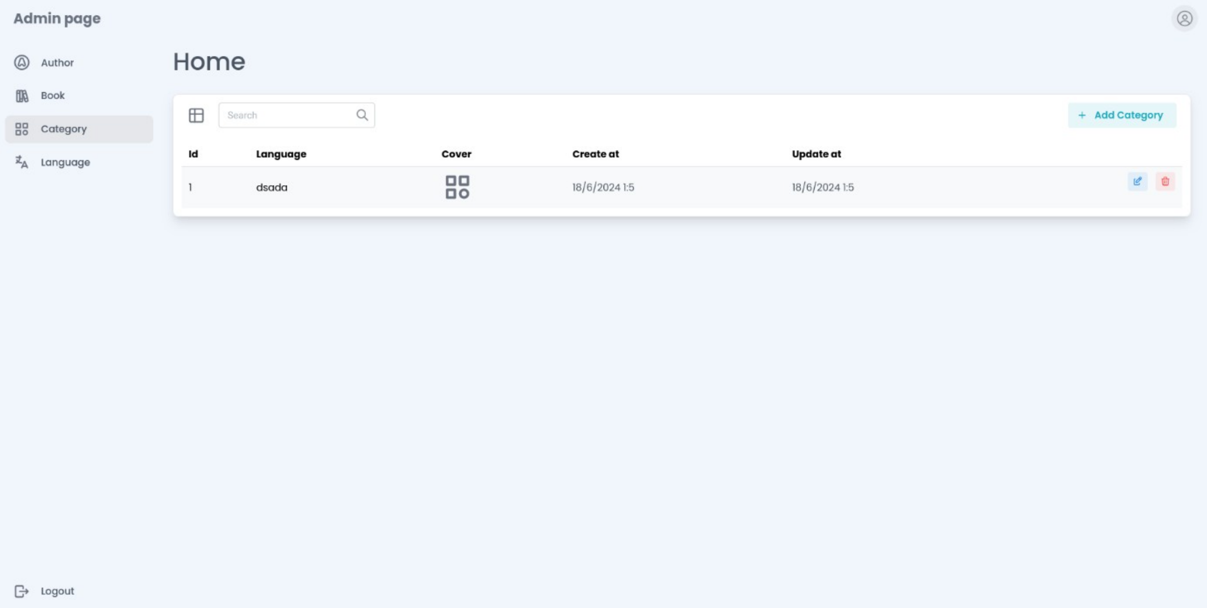
Description

Enter Author description

Cancel Create

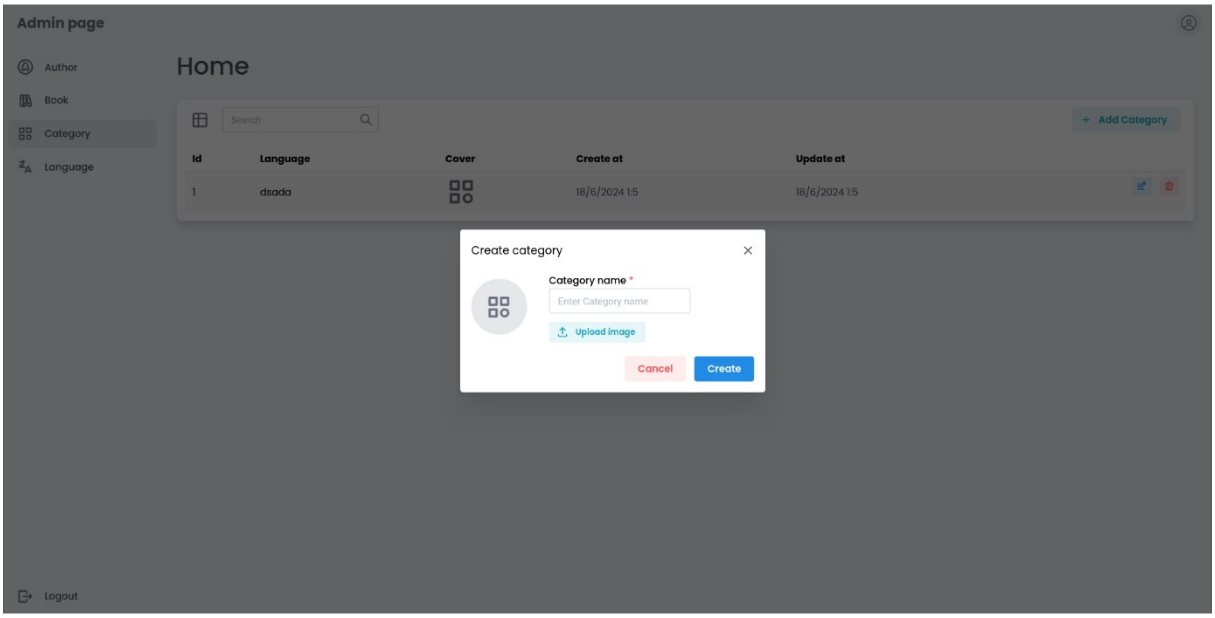
Picture 4. 3 Create author page

4.1.4 Category management page



Id	Language	Cover	Create at	Update at
1	dsada		18/6/2024 1:5	18/6/2024 1:5

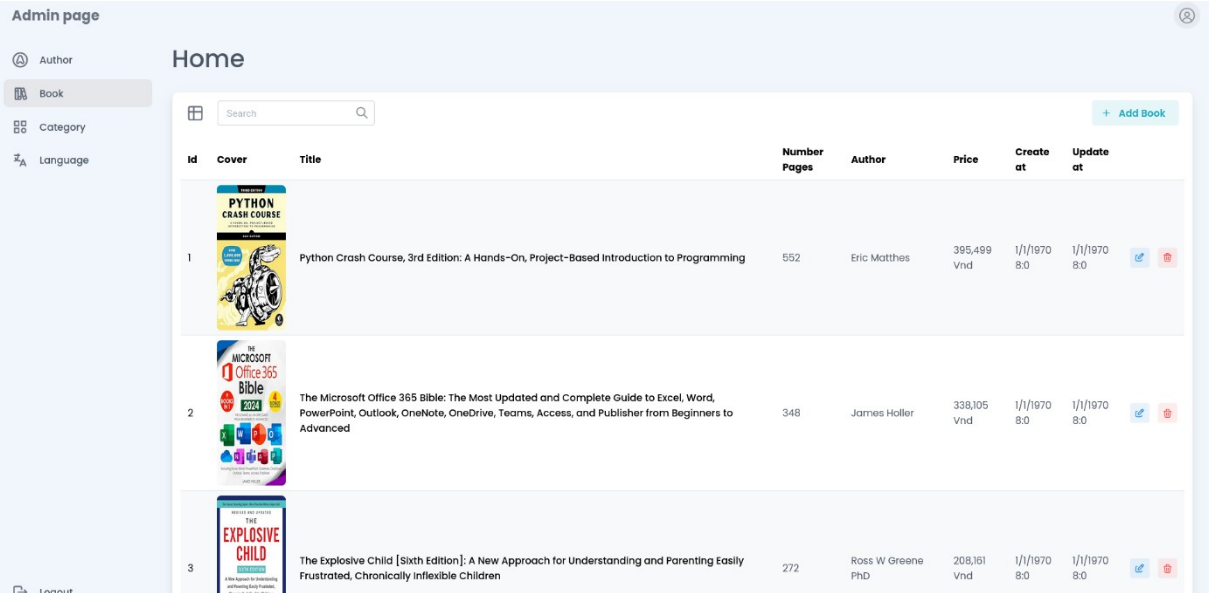
Picture 4. 4 Category management page



Id	Language	Cover	Create at	Update at
1	dsada		18/6/2024 1:5	18/6/2024 1:5

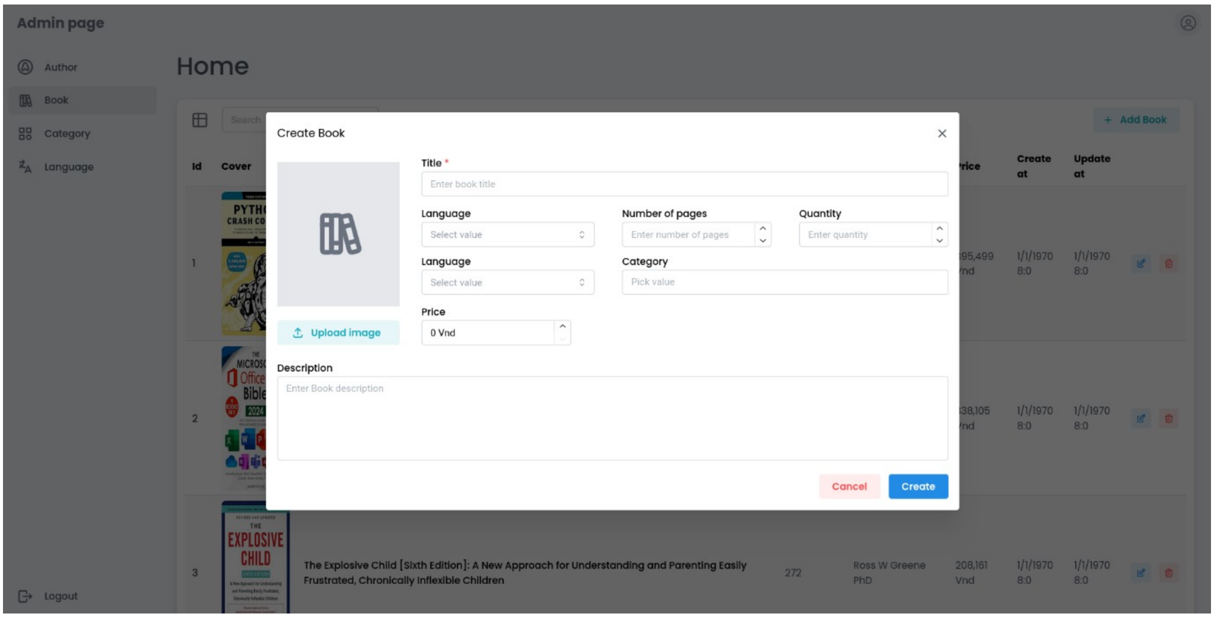
Picture 4. 5 Create category page

4.1.5 Book management page



Id	Cover	Title	Number Pages	Author	Price	Create at	Update at
1		Python Crash Course, 3rd Edition: A Hands-On, Project-Based Introduction to Programming	552	Eric Matthes	395,499 Vnd	1/1/1970 8:0	1/1/1970 8:0
2		The Microsoft Office 365 Bible: The Most Updated and Complete Guide to Excel, Word, PowerPoint, Outlook, OneNote, OneDrive, Teams, Access, and Publisher from Beginners to Advanced	348	James Holler	338,105 Vnd	1/1/1970 8:0	1/1/1970 8:0
3		The Explosive Child [Sixth Edition]: A New Approach for Understanding and Parenting Easily Frustrated, Chronically Inflexible Children	272	Ross W Greene PhD	208,161 Vnd	1/1/1970 8:0	1/1/1970 8:0

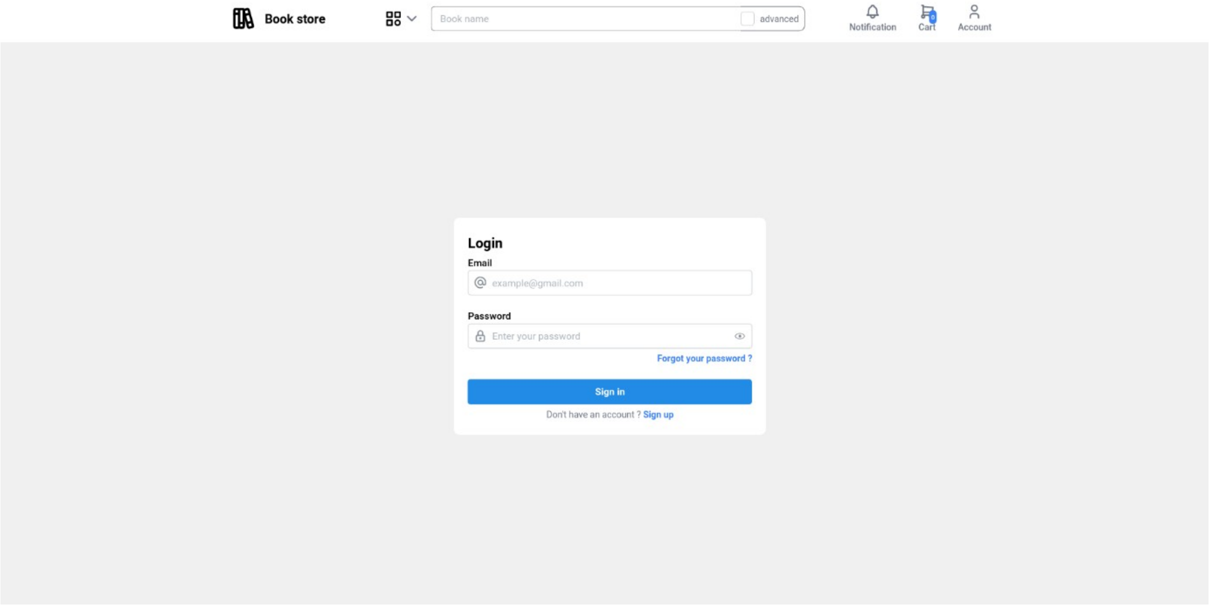
Picture 4. 6 Book management page



Picture 4. 7 Create book page

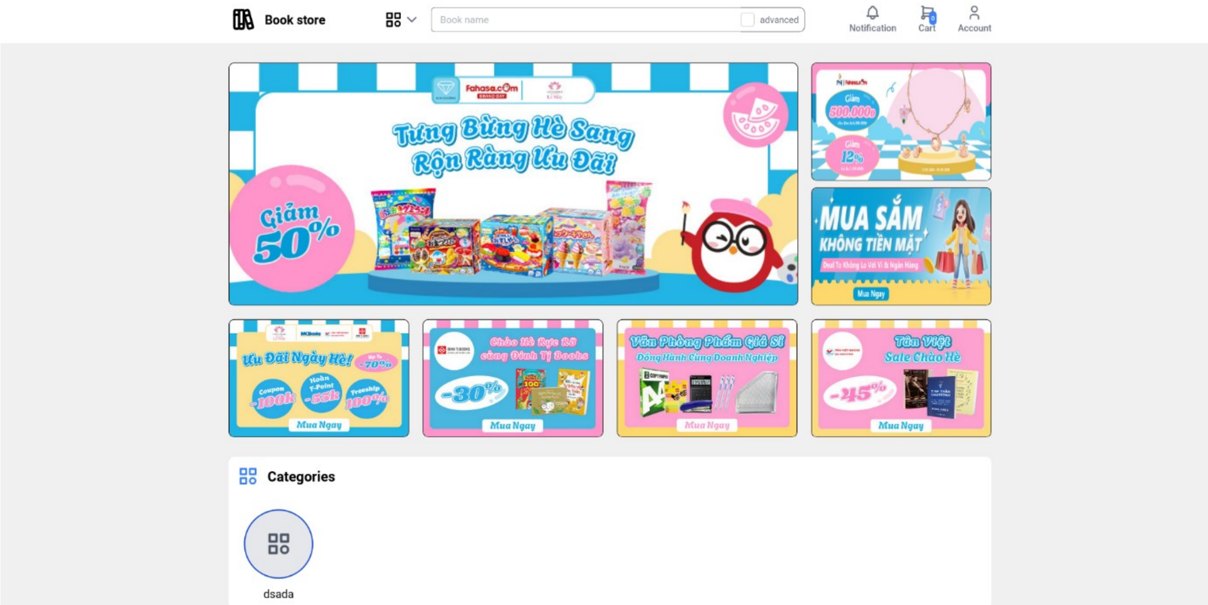
4.2 Customer page

4.2.1 Login page

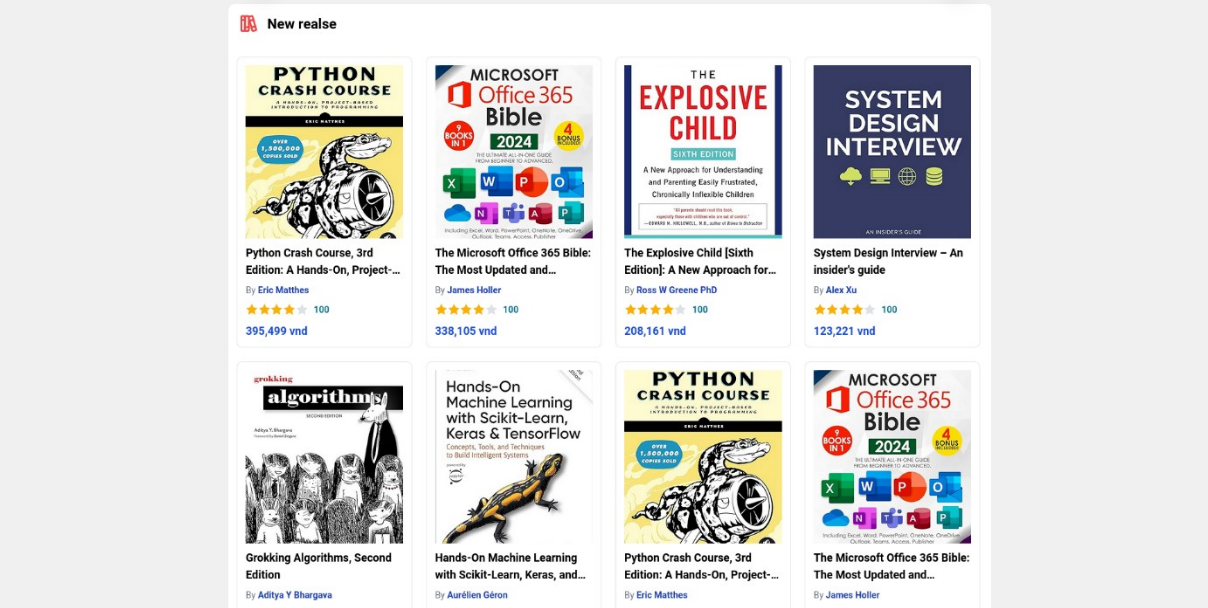


Picture 4. 8 Customer login page

4.2.2 Home page



Picture 4. 9 Home Page



Python Crash Course, 3rd Edition: A Hands-On, Project-...
By Eric Matthes
★★★★☆ 100
395,499 vnd

The Microsoft Office 365 Bible: The Most Updated and...
By James Holler
★★★★☆ 100
338,105 vnd

The Explosive Child [Sixth Edition]: A New Approach for...
By Ross W Greene PhD
★★★★☆ 100
208,161 vnd

System Design Interview – An insider's guide
By Alex Xu
★★★★☆ 100
123,221 vnd

Grokking Algorithms, Second Edition
By Aditya Y Bhargava

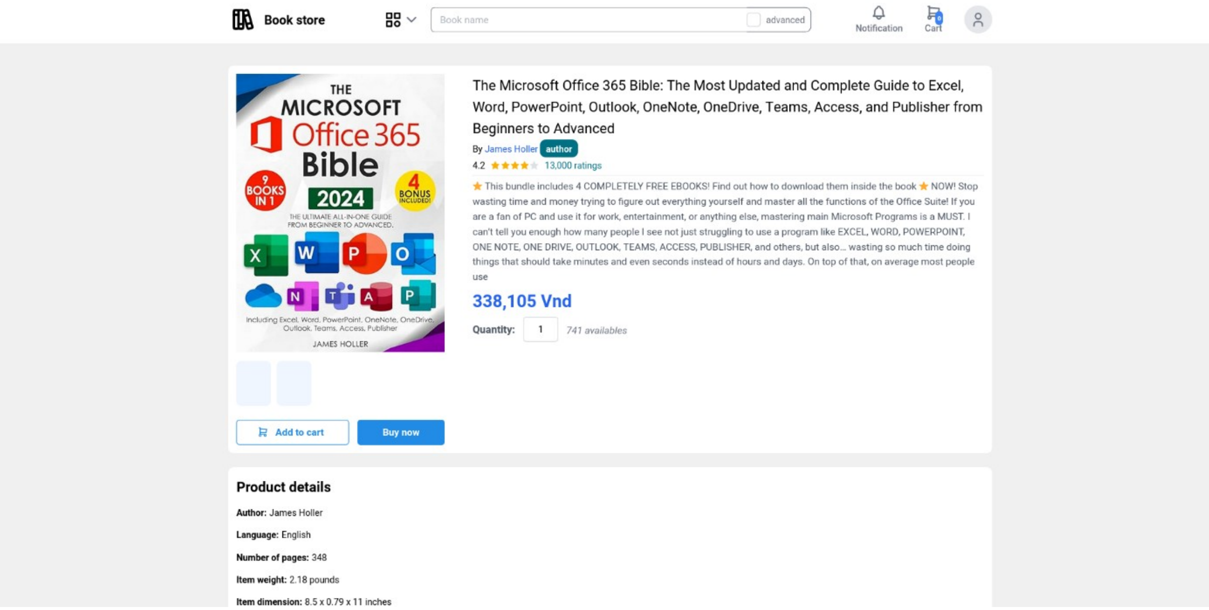
Hands-On Machine Learning with Scikit-Learn, Keras, and...
By Aurélien Géron

Python Crash Course, 3rd Edition: A Hands-On, Project-...
By Eric Matthes

The Microsoft Office 365 Bible: The Most Updated and...
By James Holler

Picture 4. 10 New release section

4.2.3 Book page


Book store
Book name
advanced
Notification
Cart
0
THE MICROSOFT Office 365 Bible
9 BOOKS IN 1
2024
4 BONUS INCLUDED!
THE ULTIMATE ALL-IN-ONE GUIDE FROM BEGINNER TO ADVANCED.
Including Excel, Word, PowerPoint, OneNote, OneDrive, Outlook, Teams, Access, Publisher
JAMES HOLLER
Add to cart
Buy now

The Microsoft Office 365 Bible: The Most Updated and Complete Guide to Excel, Word, PowerPoint, Outlook, OneNote, OneDrive, Teams, Access, and Publisher from Beginners to Advanced

By James Holler author

4.2 ★★★★☆ 13,000 ratings

★ This bundle includes 4 COMPLETELY FREE EBOOKS! Find out how to download them inside the book ★ NOW! Stop wasting time and money trying to figure out everything yourself and master all the functions of the Office Suite! If you are a fan of PC and use it for work, entertainment, or anything else, mastering main Microsoft Programs is a MUST. I can't tell you enough how many people I see not just struggling to use a program like EXCEL, WORD, POWERPOINT, ONE NOTE, ONE DRIVE, OUTLOOK, TEAMS, ACCESS, PUBLISHER, and others, but also... wasting so much time doing things that should take minutes and even seconds instead of hours and days. On top of that, on average most people use

338,105 Vnd

Quantity: 1 741 availables

Product details

Author: James Holler

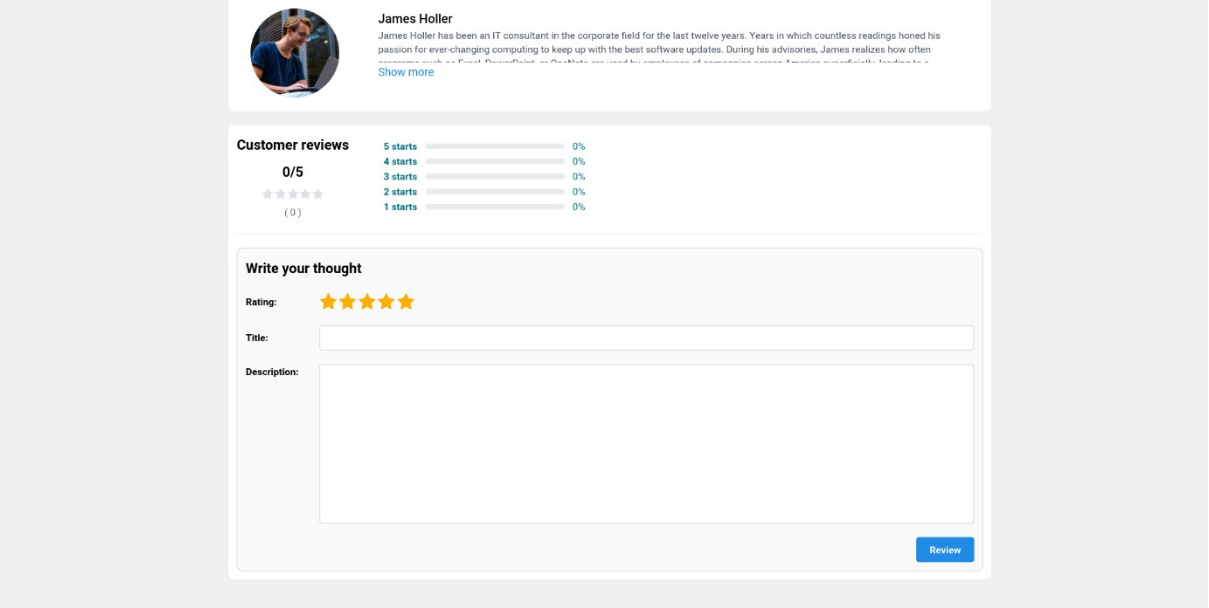
Language: English

Number of pages: 348

Item weight: 2.18 pounds

Item dimension: 8.5 x 0.79 x 11 inches

Picture 4. 11 Book detail page



James Holler

James Holler has been an IT consultant in the corporate field for the last twelve years. Years in which countless readings honed his passion for ever-changing computing to keep up with the best software updates. During his advisories, James realizes how often programs such as Excel, PowerPoint, or OneNote are used by employees of companies across America superficially, leading to a

Customer reviews

0/5

☆☆☆☆☆

(0)

5 starts	0%
4 starts	0%
3 starts	0%
2 starts	0%
1 starts	0%

Write your thought

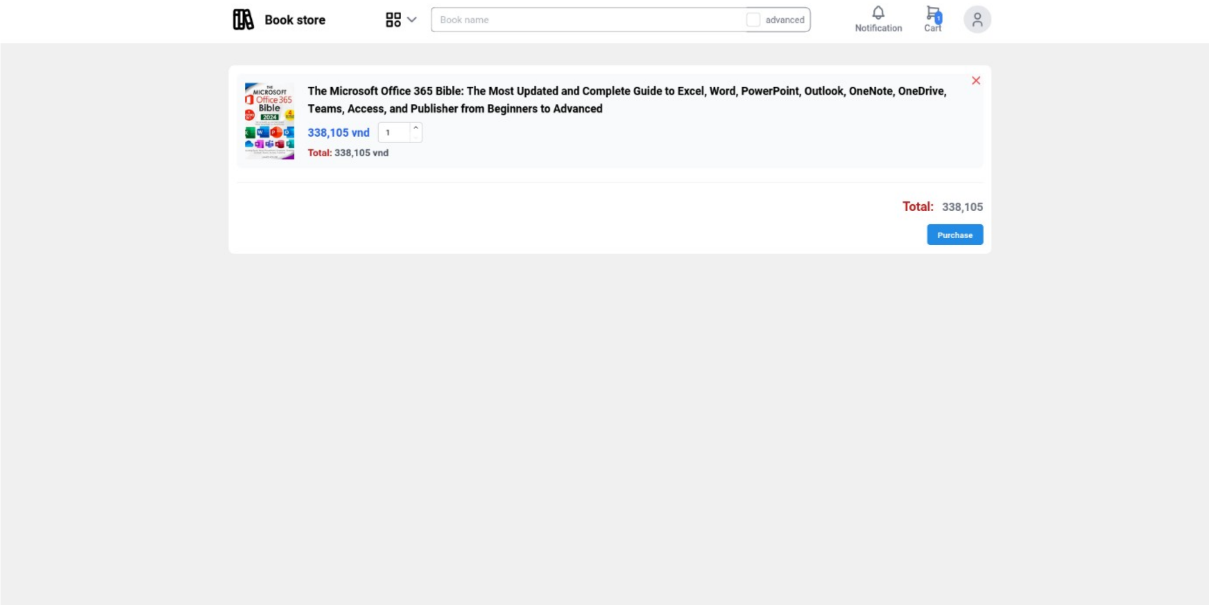
Rating: ★★★★★

Title:

Description:

Picture 4. 12 Review book section

4.2.4 Cart page



The Microsoft Office 365 Bible: The Most Updated and Complete Guide to Excel, Word, PowerPoint, Outlook, OneNote, OneDrive, Teams, Access, and Publisher from Beginners to Advanced

338,105 vnd 1

Total: 338,105 vnd

Total: 338,105

Picture 4. 13 Cart page

4.3 Building collaborative filtering model

4.3.1. Dataset overview

- Movielens dataset: is a widely-used collection of movie ratings and user data, provided by the GroupLens Research Project at the University of Minnesota. It is primarily used for research in collaborative filtering, recommendation systems, and machine learning. Dataset include:

User Ratings:

- The dataset includes millions of ratings (on a 5-star scale) from thousands of users.
- Each rating is associated with a specific user and movie, with timestamps indicating when the rating was given.

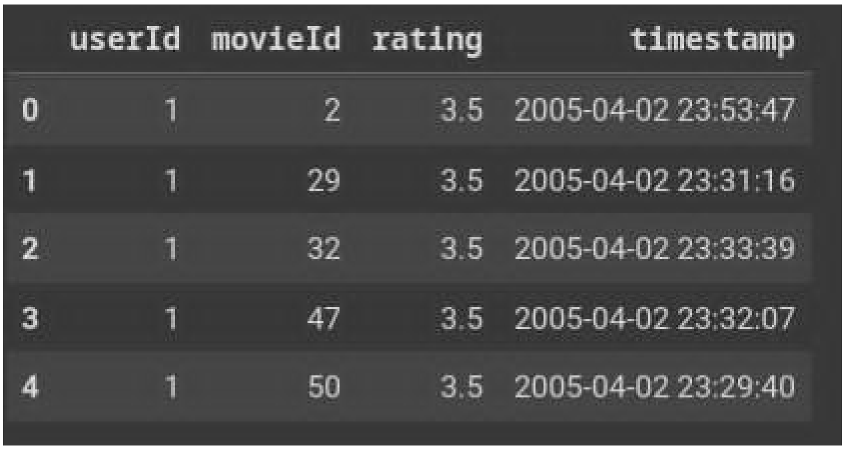
Movies:

- Each movie in the dataset is identified by a unique movie ID.
- Metadata for movies includes titles, genres, and release dates.

Users:

- Users are anonymized with unique user IDs.

Even Dataset provide a lot of information, but we only need user rating data using for training and evaluation neural collaborative filtering model. So that we don't care other data than user ratings

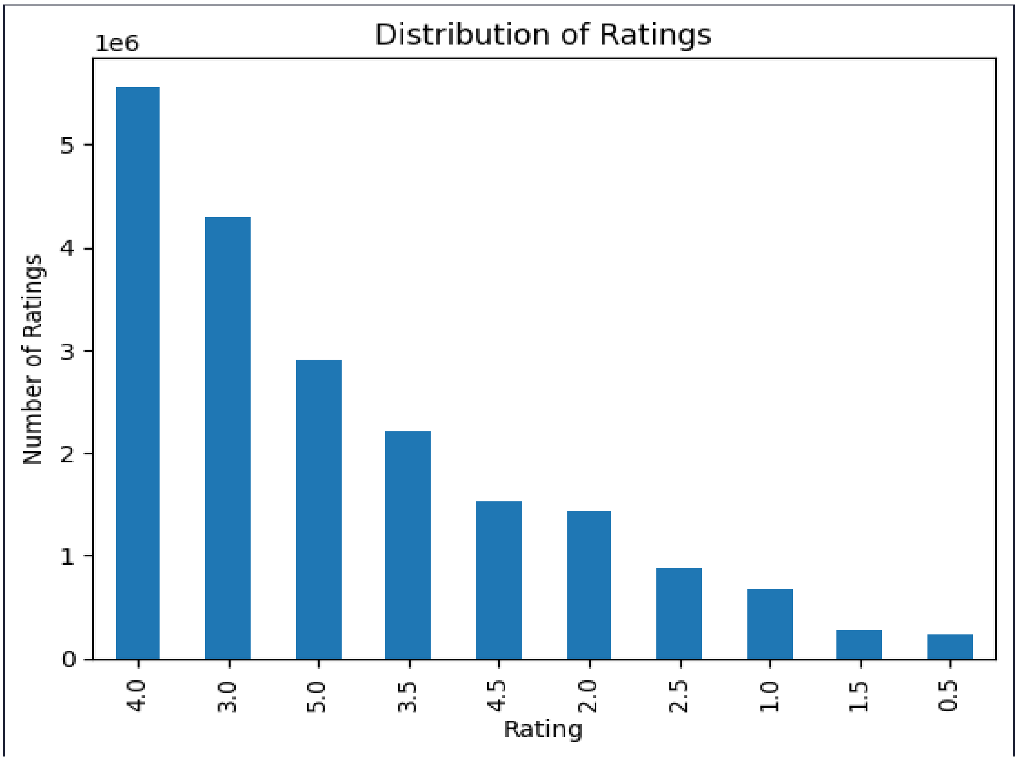


	userId	movieId	rating	timestamp
0	1	2	3.5	2005-04-02 23:53:47
1	1	29	3.5	2005-04-02 23:31:16
2	1	32	3.5	2005-04-02 23:33:39
3	1	47	3.5	2005-04-02 23:32:07
4	1	50	3.5	2005-04-02 23:29:40

Picture 4. 14 Rating dataset

4.3.2 Data exploration

- How rating distribute in dataset ?

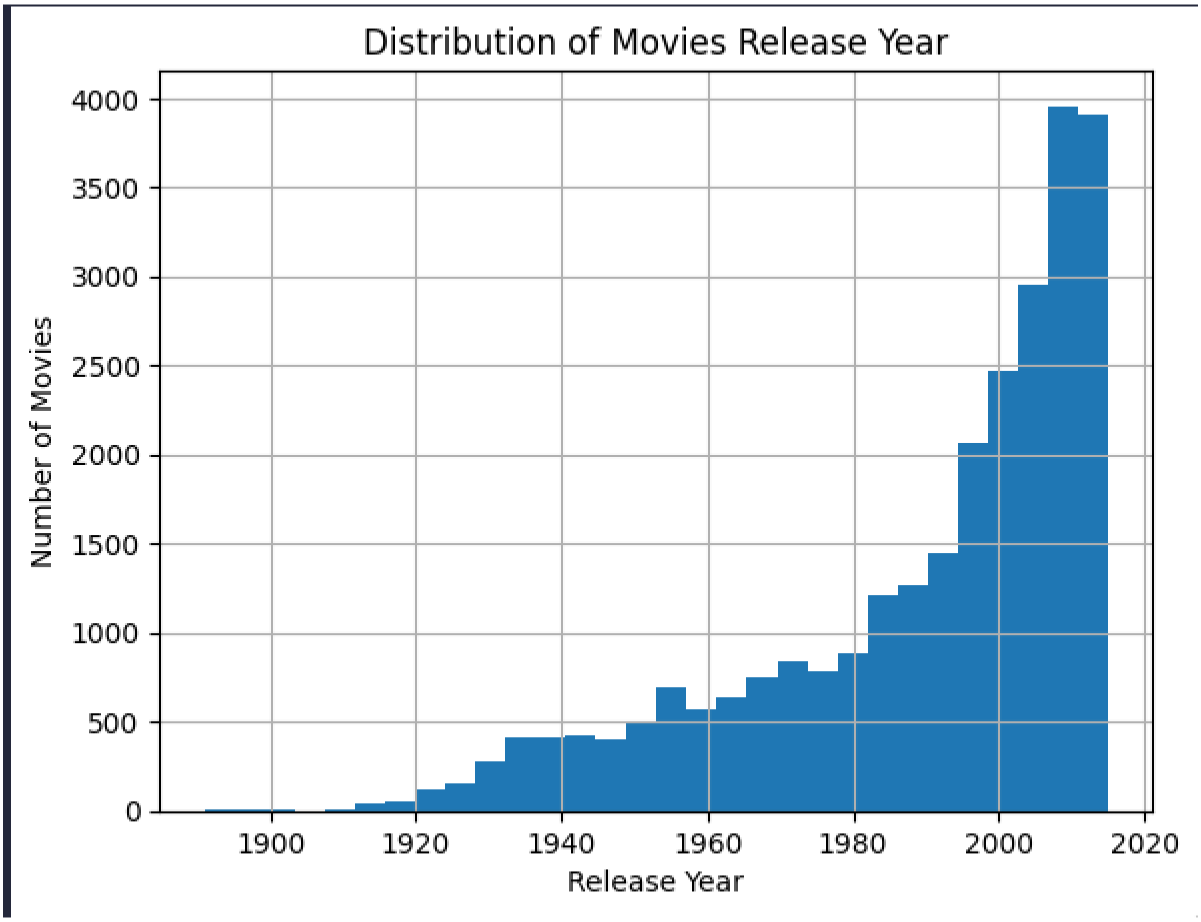


Picture 4. 15 Distribution of ratings

Analysis:

- The ratings range from 0.5 to 5.0
- The most frequent rating is 4.0, with approximately 5,000,000 ratings, which is about 16% higher than the second most frequent rating of 3.0 (4,200,000 ratings)
- In contrast, the least frequent rating is 0.5, with approximately 400,000 ratings, which is 92% lower than the most frequent rating.

- What is the distribution of movies by release year ?

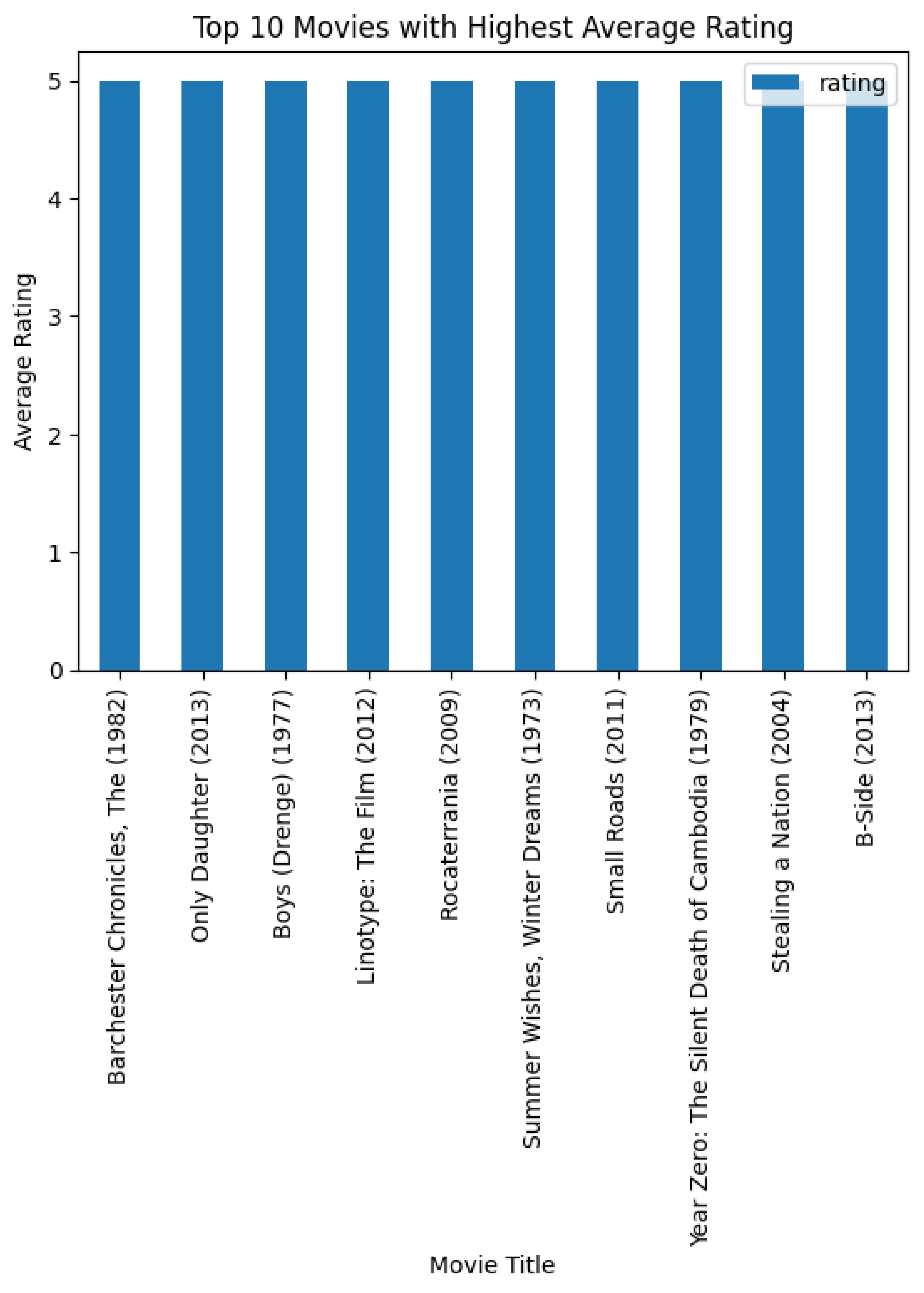


Picture 4. 16 Distribution of movies release year

Analysis:

- The chart shows that the number of movies released increases as the years progress from 1900 to 2018.
- The year with the fewest movie releases is around 1900, with only approximately 10 movies released.
- The year with the most movie releases is around 2018, with approximately 3,900 movies released.

- What is Top 10 movies with highest average rating ?

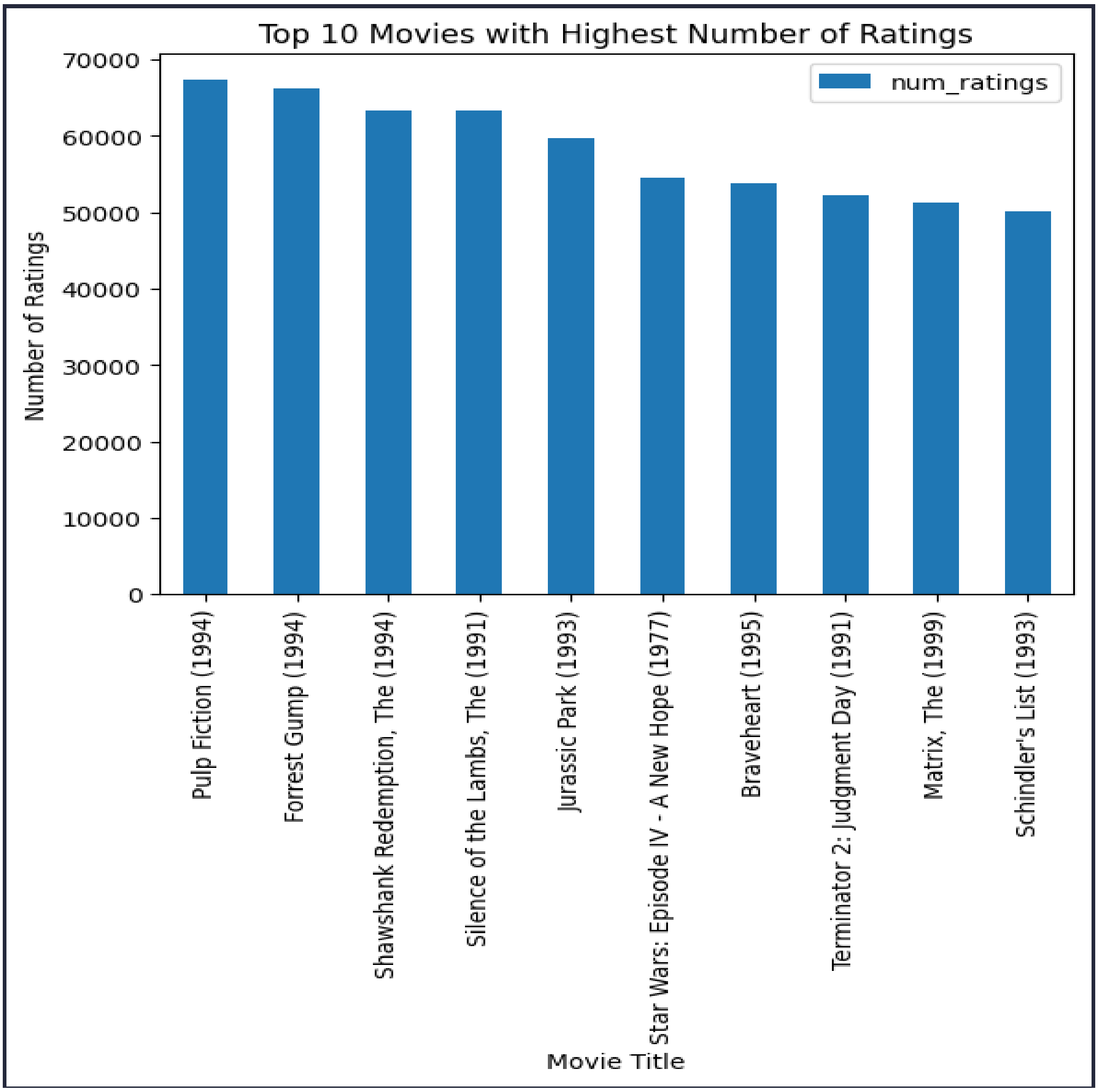


Picture 4. 17 Top 10 Movies with highest avarage rating

Analysis:

- The chart shows that all top 10 movies have an average rating of 5.0. However, These movies might have a low number of ratings, but each rating is the highest possible (5.0). This demonstrates that even if a movie has a high average rating, it does not necessarily mean the movie is popular.

- What is top 10 movies with highest number of ratings ?



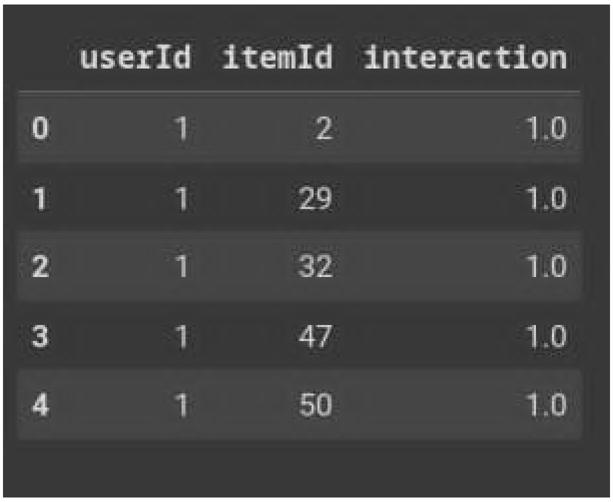
Picture 4. 18 Top 10 movies with highest number of ratings

Analysis*:*

- The movie with the highest number of ratings is "Pulp Fiction" (1994), with approximately 6,900 ratings.
- There are three movies with the highest number of ratings that were released in 1994. This indicates that in 1994, people were particularly likely to rate the movies they watched.

4.3.3. Preprocess data

Since we only need to know which movies a user has interacted with, we should convert the rating value into 1 to indicate that the user has interacted with the movie.



	userId	itemId	interaction
0	1	2	1.0
1	1	29	1.0
2	1	32	1.0
3	1	47	1.0
4	1	50	1.0

Picture 4. 19 Data after preprocess

4.3.4. Split train and test data

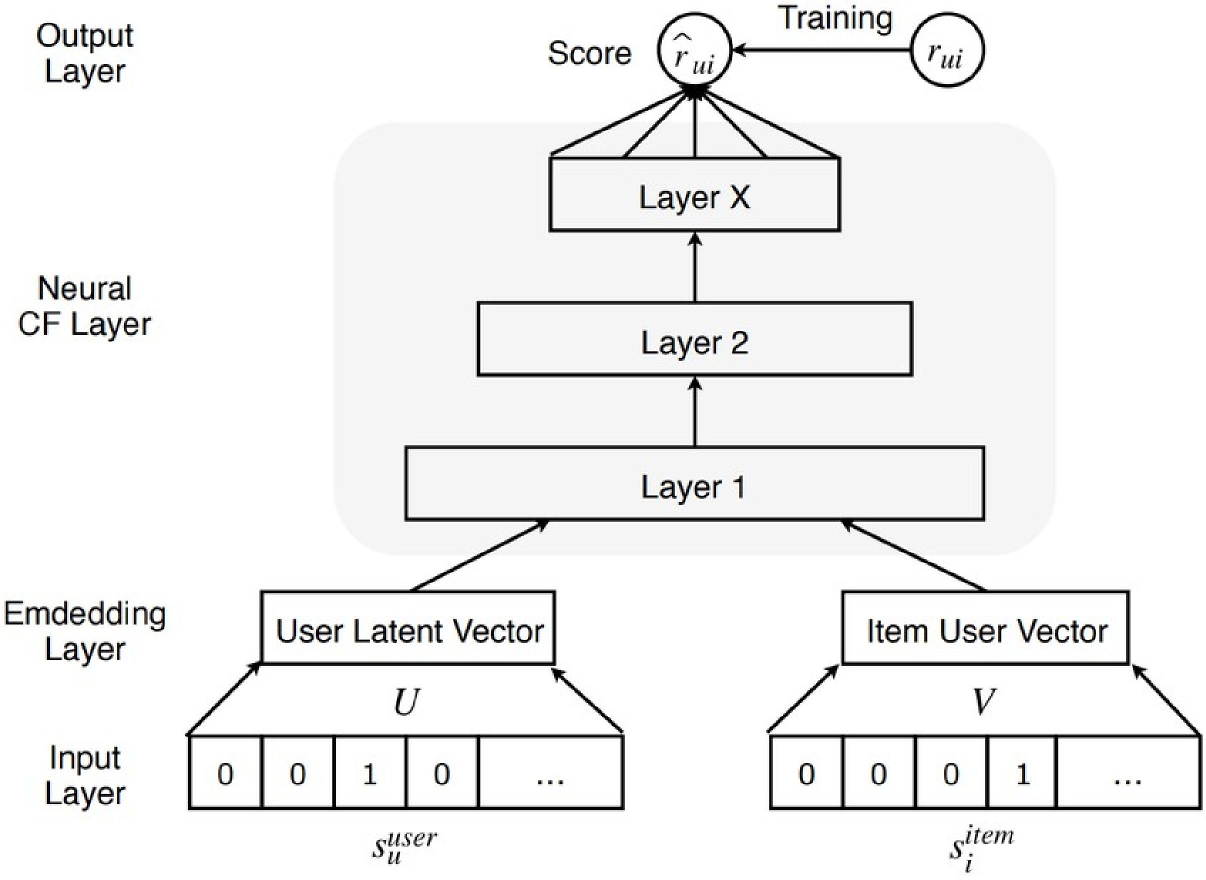
We will split the dataset into training and testing sets with an 80:20 ratio. This means 80% of the data will be used for training the model, and the remaining 20% will be used for testing its performance.

To ensure the reliability and robustness of our model, we will use a temporal split strategy. This involves training the model on historical data and evaluating it on the most recent data. By doing this, we simulate a real-world scenario where the model is trained on past interactions and tested on how well it predicts future interactions.

Furthermore, to construct a balanced dataset for training, we will create one positive interaction and four negative interactions for each user. Positive interactions are those where the user has interacted with the movie (indicated by a rating), and negative interactions are where the user has not interacted with the movie. This approach helps in addressing class imbalance and ensures that the model learns to distinguish between movies that users are likely to interact with and those they are not.

4.3.5. Building model

We use tensorflow and keras to create a neural collaborative filter



Picture 4. 20 Neural collaborative filtering architecture

The model architecture include:

- **Input layer:** The input layer typically consists of two inputs: one for the user ID and one for the item ID. These IDs are used to look up the corresponding embeddings.

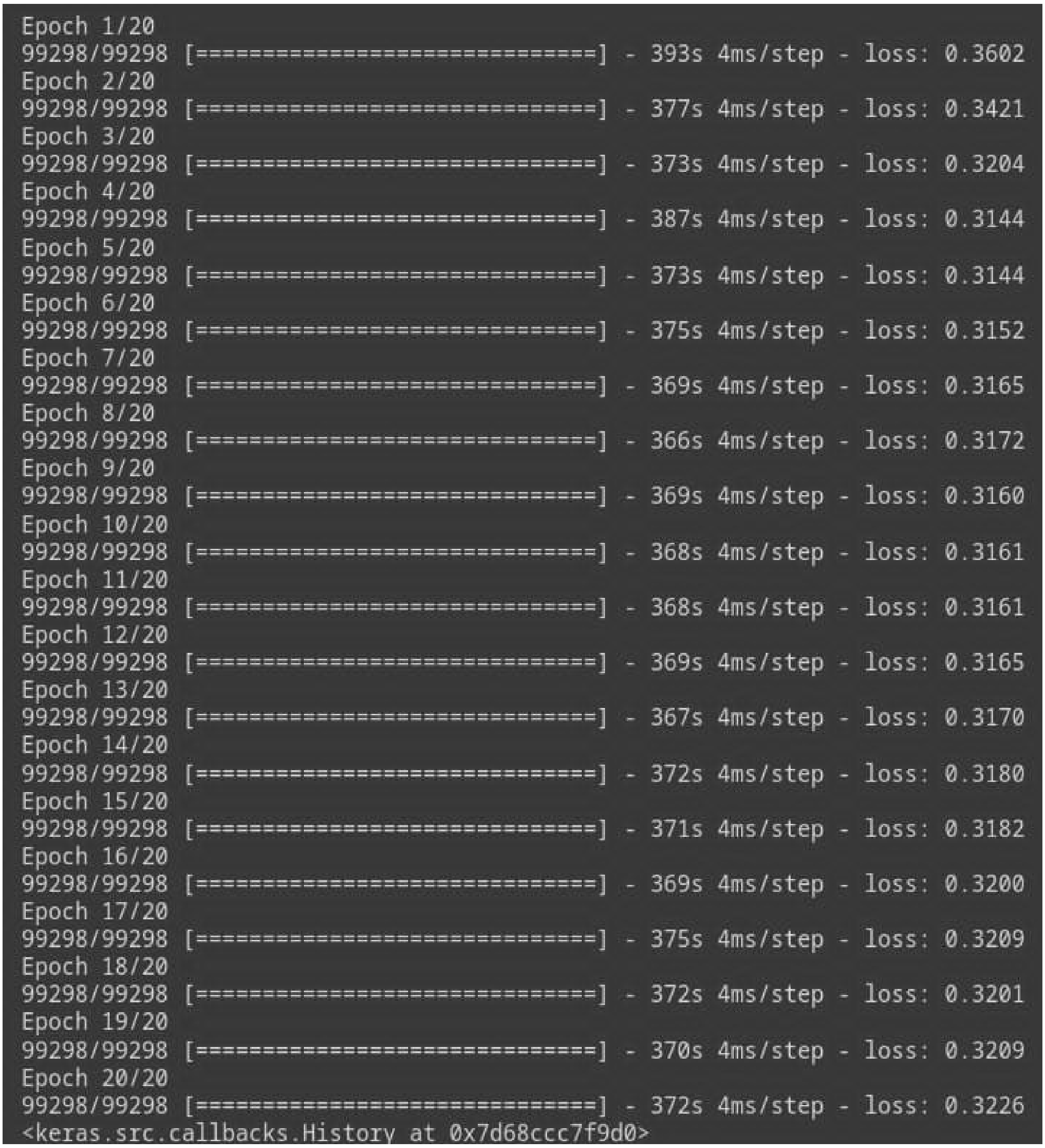
- **Embedding Layer:** These layers transform the user and item IDs into their corresponding embeddings. This transformation is often done through embedding lookup tables.

- **Neural CF Layer**: This component consists of several fully connected layers with activation functions (like ReLU). It takes the concatenated embeddings (or the result of another interaction mechanism) as input and learns a non-linear function to predict the interaction between users and items.

- **Output layer**: The output layer consists of a single neuron with a sigmoid activation function to predict the probability of interaction between a user and an item.

4.3.6.Training model

We will train dataset with 20 epoch



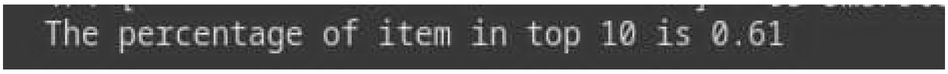
```
Epoch 1/20
99298/99298 [==============================] - 393s 4ms/step - loss: 0.3602
Epoch 2/20
99298/99298 [==============================] - 377s 4ms/step - loss: 0.3421
Epoch 3/20
99298/99298 [==============================] - 373s 4ms/step - loss: 0.3204
Epoch 4/20
99298/99298 [==============================] - 387s 4ms/step - loss: 0.3144
Epoch 5/20
99298/99298 [==============================] - 373s 4ms/step - loss: 0.3144
Epoch 6/20
99298/99298 [==============================] - 375s 4ms/step - loss: 0.3152
Epoch 7/20
99298/99298 [==============================] - 369s 4ms/step - loss: 0.3165
Epoch 8/20
99298/99298 [==============================] - 366s 4ms/step - loss: 0.3172
Epoch 9/20
99298/99298 [==============================] - 369s 4ms/step - loss: 0.3160
Epoch 10/20
99298/99298 [==============================] - 368s 4ms/step - loss: 0.3161
Epoch 11/20
99298/99298 [==============================] - 368s 4ms/step - loss: 0.3161
Epoch 12/20
99298/99298 [==============================] - 369s 4ms/step - loss: 0.3165
Epoch 13/20
99298/99298 [==============================] - 367s 4ms/step - loss: 0.3170
Epoch 14/20
99298/99298 [==============================] - 372s 4ms/step - loss: 0.3180
Epoch 15/20
99298/99298 [==============================] - 371s 4ms/step - loss: 0.3182
Epoch 16/20
99298/99298 [==============================] - 369s 4ms/step - loss: 0.3200
Epoch 17/20
99298/99298 [==============================] - 375s 4ms/step - loss: 0.3209
Epoch 18/20
99298/99298 [==============================] - 372s 4ms/step - loss: 0.3201
Epoch 19/20
99298/99298 [==============================] - 370s 4ms/step - loss: 0.3209
Epoch 20/20
99298/99298 [==============================] - 372s 4ms/step - loss: 0.3226
<keras.src.callbacks.History at 0x7d68ccc7f9d0>
```

Picture 4. 21 Training model

4.3.7. Evaluate model

Metrics: When evaluating the performance of recommendation models, it is often more insightful to use ranking-based metrics instead of simple accuracy. One effective metric for this purpose is the Hit Rate (HR). This metric assesses how often the true positive item (e.g., a movie the user actually liked) appears in the top-K recommendations made by the model.

Result:



```
The percentage of item in top 10 is 0.61
```

Picture 4. 22 Evaluation model metric results

4.3.7 Conclusion

The model's prediction accuracy is approximately 61%, which is not satisfactory. This performance could be attributed to several factors:

- The dataset might not be sufficiently large, which can limit the model's ability to learn and generalize effectively

- The model architecture itself might not be well-suited for this particular problem. It’s crucial to select a model that aligns well with the nature of the data and the task at hand.

4.4 Evaluation

Advantage:

- Application provides enough features for managing books
- Application integrates advance deep learning model to expand application features and functionality

Disadvantage:

- UI/UX is not good, we need more time to refactor and improve the User interface and user experience
- Deep learning model is not good to solve the problem for recommender system

CONCLUSION

1. Archived result:

During the research and implementation phase of the project, the following results were achieved:

- Understand how process of software development work
- Using All taught knowledge to apply to project
- Understand how to use popular libraries like Laravel, ReactJS, ...
- Understand how recommender system work
- Understand how to integrate deep learning model into application

2. Needed improvement

- Need improve design UI/UX to increase user experience
- Need to find way to optimize recommendation system model when there are large users
- Add more features
- Need to improve performance of recommender systems deep learning model

REFERENCES

[1] Laravel, URL: https://laravel.com/

[2] ReactJS, URL: https://react.dev/

[3] BERT, URL: https://www.geeksforgeeks.org/explanation-of-bert-model-nlp/

[4] Neural collaboration filtering, URL: https://medium.com/data-science-in-your-pocket/recommendation-systems-using-neural-collaborative-filtering-ncf-explained-with-codes-21a97e48a2f7

DA NANG UNIVERSITY

UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY......................................

SOCIALIST REPUBLIC OF VIETNAM

Independence - Freedom - Happiness

GRADUATION PROJECT COMMENT

- **General information:**
- Student name: …………………………………………………………………..
- Class: …………………………………… Student ID ………………………….
- Topic title: ……………………………………………………………………
- Instructor: ……………………… Academic title/ degree: ………..........
- **Reviews of graduation project**
- About the urgency, novelty, usability of the topic: (2 points)

  ……………………………………………………………………………

  ……………………………………………………………………………
- About the results of solving the tasks required by the project: (4 points)

  ……………………………………………………………………………

  ……………………………………………………………………………
- About the form, structure and layout of the graduation project: (2 points)

  ……………………………………………………………………………

  ……………………………………………………………………………
- The topic includes scientific value/article/problem solving of the enterprise or school: (1 point)

  ……………………………………………………………………………

  ……………………………………………………………………………
- Existing shortcoming need to be supplemented or modified:

  ……………………………………………………………………………

  ……………………………………………………………………………
- **Spirit and attitude of the student (1 point):**

  ……………………………………………………………………………

  ……………………………………………………………………………
- **Evaluation:**
- Evaluation point: …./10
- Suggest: Defense permitted/ Edit to defend/ Defense not permitted

Da Nang, date……month……. 2024

Instructor