Independent University **Bangladesh (IUB)**

IUB Academic Repository

Internship Reports Summer 2023

2023-10-19

An Undergraduate Internship/Project on Book Recommendation System at ADN DigiNet

Kabir, Ashraful

Independent University, Bangladesh

https://ar.iub.edu.bd/handle/11348/703

Downloaded from IUB Academic Repository



An Undergraduate Internship/Project on Book Recommendation System at ADN DigiNet

By

Ashraful Kabir

Student ID: **2022700**

Summer, 2023

Supervisor:

Marzan Binte Hassan

Adjunct Lecturer

Department of Computer Science & Engineering

Independent University, Bangladesh

October 19, 2023

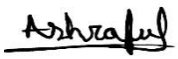
Dissertation submitted in partial fulfillment for the degree of
Bachelor of Science in Computer Science

Department of Computer Science & Engineering

Independent University, Bangladesh

Attestation

This attestation serves to confirm that the report entitled "**An Undergraduate Internship/Project on Book Recommendation System at ADN DigiNet**" has been executed by me, Ashraful Kabir (2022700), and submitted as a component of the requirements for the Degree of Computer Science from Independent University, Bangladesh (IUB). The report was completed under the guidance of Marzan Binte Hasan (Supervisor). I further attest that all the work presented in this report is entirely original and was acquired during my tenure as an intern. All sources of information utilized in the project and report have been duly acknowledged and cited in accordance with academic standards.



19/10/2023

Signature

Date

Ashraful Kabir

Name

Acknowledgement

Special thanks to the Machine Learning trainers at ADN DigiNet for doing the following:

- Onboarding me (the intern) into their workplace
- Teaching me the basics of artificial intelligence, machine learning, deep learning and guiding me to learn further
- Interacting with me just like another team member
- Grooming and training me with relevant skills, making me a potential future candidate for becoming a permanent team member

Letter of Transmittal

Marzan Binte Hassan
Adjunct Lecturer
Department of Computer Science and Engineering
School of Engineering and Computer Science
Independent University, Bangladesh

Subject: **Submission of Final Internship Report in Compliance with Graduation Requirements.**

Dear Ma'am,

With due respect, I am submitting my Internship Report, which is a requisite component of my Bachelor's degree in Computer Science and Engineering. It is a profound honor to have had the opportunity to work under your esteemed guidance. This report is a result of my internship tenure at ADN DigiNet, where I was privileged to work for a period of three months. This internship afforded me the opportunity to gain both theoretical and practical knowledge, as well as to establish professional connections. I have made a conscientious effort to ensure that this report is as informative and comprehensive as possible, by incorporating the skills and knowledge that I acquired during my internship. I have adhered to the guidelines provided and have furnished sufficient detail in the report. I am convinced that this report will satisfactorily fulfill the purpose of my internship program.

I would be deeply appreciative if you would kindly review the report and offer your invaluable feedback. It would be a great honor if you found the report to be informative and of use in gaining a clearer perspective on my internship experience.

Sincerely,

Ashraful Kabir
ID: 2022700
Department of Computer Science and Engineering
Independent University, Bangladesh

Evaluation Committee

Supervision Panel

..... Academic Supervisor Name: Marzan Binte Hasan	..... Industry Supervisor Name: Dodul Haque Khan

Panel Members

..... Panel Member-1 Name: Sanzar Adnan Alam	..... Panel Member-2 Name: Mohammad Motiur Rahman

Panel Members

..... Panel Member-3 Name: Md. Mahmudul Peyal	..... Panel Member-4 Name: Mahir Al Kamal

Office Use

..... Program Coordinator Name: Subrata Kumar Dey	..... Head of Department Computer Science & Engineering Name: Mahady Hasan

Abstract

As an intern in the artificial intelligence division of ADN DigiNet, I had the opportunity of developing prototypes and proof of concept projects utilizing Machine Learning (ML) and Artificial Intelligence (AI) techniques. The internship commenced with an intensive two-month training phase, immersing us in ML concepts, tools, and available libraries. Subsequently, the final month was dedicated to applying our newly acquired knowledge and skills to work on individual projects.

The development of an intelligent book recommendation system was the main goal of my project, that utilizes AI techniques to suggest popular books to users and recommend books based on similarity. The primary objective of the system is to address the growing challenge of information overload in the digital age, where countless books are published regularly, making it increasingly difficult for users to discover relevant and interesting reads.

As part of this project, I played a key role in developing the system's AI model. I was specifically in charge of putting the essential features into place, such as the collaborative and popularity-based filtering methods. Throughout the development process, various challenges were addressed, including data preprocessing and model optimization. By employing suitable AI techniques, the model successfully overcomes these challenges, resulting in an effective book recommendation system.

Table of Contents

Attestation 1

Acknowledgement 2

Letter of Transmittal 3

Evaluation Committee 4

Abstract 5

Introduction 10

1.1 Overview/Background of the Work 10
1.2 Objectives 10
1.3 Scopes 10

Literature Review 11

2.1 Relationship with Undergraduate Studies 11
2.2 Related works 11

Project Management & Financing 12

3.1 Work Breakdown Structure 12
3.2 Process/Activity wise Time Distribution 12
3.3 Gantt Chart 13
3.4 Process/Activity wise Resource Allocation 13
3.5 Estimated Costing 13

Methodology 14

Body of the Project 15

5.1 Work Description 15
5.2 Requirement Analysis 16
5.3 System Analysis 18
5.3.1 Six Element Analysis 18
5.3.2 Feasibility Analysis 23
5.3.3 Problem Solution Analysis 23
5.3.4 Effect and Constraints Analysis 24
5.4 System Design 24
5.5 Implementation 25
5.6 Testing 26

Results & Analysis 28

Project as Engineering Problem Analysis 30

7.1 Sustainability of the Project/Work 30

7.2 Social and Environmental Effects and Analysis 30
7.3 Addressing Ethics and Ethical Issues 30
Lesson Learned 31
8.1 Challenges Faced During this Period 31
Future Work & Conclusion 32
9.1 Future Works 32
9.2 Conclusion 33
Bibliography 34

List of Figures

Figure 1: Timeline of our project 13
Figure 2: Rich Picture of the system 16
Figure 3: Entity Relationship Diagram (ERD) of the system 24
Figure 4: Data of top 50 popular books 26
Figure 5: Output of recommend function when "1984" is send as parameter 27
Figure 6: Output of recommend function when “A Bend in the Road” is send as parameter 27
Figure 7: Displaying the popular books in UI 28
Figure 8: Displaying the similar books when a book name is given as input in UI 29

List of Tables

Table 1: Six Element Analysis of the system 18

Chapter 1

Introduction

1.1 Overview/Background of the Work

The book industry has seen exponential growth in recent years, with an overwhelming number of books being published across various genres. As a result, readers often struggle to find books that align with their preferences and interests, leading to a suboptimal reading experience. Additionally, traditional bookstores and online platforms may not always provide personalized recommendations, making it challenging for readers to discover new and engaging content. The challenge at hand was to build and develop a system for personalized book recommendations that uses machine learning (ML) models to deliver recommendations based on user preferences and historical data in order to meet these problems.

1.2 Objectives

The objective of the project was defined as developing an AI-based book recommendation system that utilizes popularity-based and collaborative filtering techniques to suggest relevant and in demand book choices to users. The system leverages user and book data attributes to generate accurate and diverse book suggestions.

1.3 Scopes

- Evaluate system performance and gather user feedback for improvement
- Ensure scalability, address challenges, and document the project comprehensively

Chapter 2

Literature Review

2.1 Relationship with Undergraduate Studies

The field of Artificial Intelligence (AI) and machine learning has seen rapid growth and has found numerous applications in various domains, including recommendation systems. AI-based recommendation systems have become integral in enhancing user experiences and increasing user engagement on online platforms. In the context of undergraduate studies, incorporating AI and machine learning techniques into educational settings has the potential to revolutionize the learning process and improve students' academic performance. The Numerical Methods, Artificial Intelligence, Machine Learning and Image Processing courses of IUB were aligned with my work.

2.2 Related works

In recent years, Bangladeshi companies have been increasingly exploring the adoption of AI and machine learning technologies to improve their operations, enhance customer experiences, and drive business growth. While various industries are leveraging AI, the following works highlight some notable applications of AI in the context of Bangladeshi companies:

- **"AI-Driven Book Recommendations for Bangladeshi Online Bookstores"** by Ahmed et al. (2022): Online bookstores in Bangladesh are leveraging AI-powered recommendation systems to provide personalized book suggestions to customers. By analyzing reading preferences, author interests, and genre choices, the model enhances the book discovery experience and encourages book sales for Bangladeshi bookstores.
- **"AI-Enhanced Financial Fraud Detection for Bangladeshi Banks"** by Khan et al. (2023): Financial fraud is a persistent concern for banks in Bangladesh. This research focuses on implementing AI-based fraud detection systems that employ anomaly detection

algorithms to identify suspicious transactions and activities. The model helps in preventing fraudulent transactions and safeguarding the financial interests of customers and banks alike.

These related works demonstrate the increasing interest of Bangladeshi companies in harnessing the potential of AI to address specific challenges and improve their operations.

Chapter 3

Project Management & Financing

3.1 Work Breakdown Structure

The timelines, resource allocation, and project milestones were all defined during a thorough planning phase. Project management tools (GitHub) were used to track the progress of the smaller tasks that made up the project. The project was divided into frontend, backend, feature engineering, data pretreatment, collecting, and algorithm implementation for recommendations. Frequent coordination meetings fostered a coherent development process by ensuring that all team members were informed of each other's progress and issues.

3.2 Process/Activity wise Time Distribution

The team set clear benchmarks to monitor progress. The project was broken up into several 2-to-5-day sprints. This strategy made it possible to complete important deliverables on schedule and with a methodical development process.

3.3 Gantt Chart

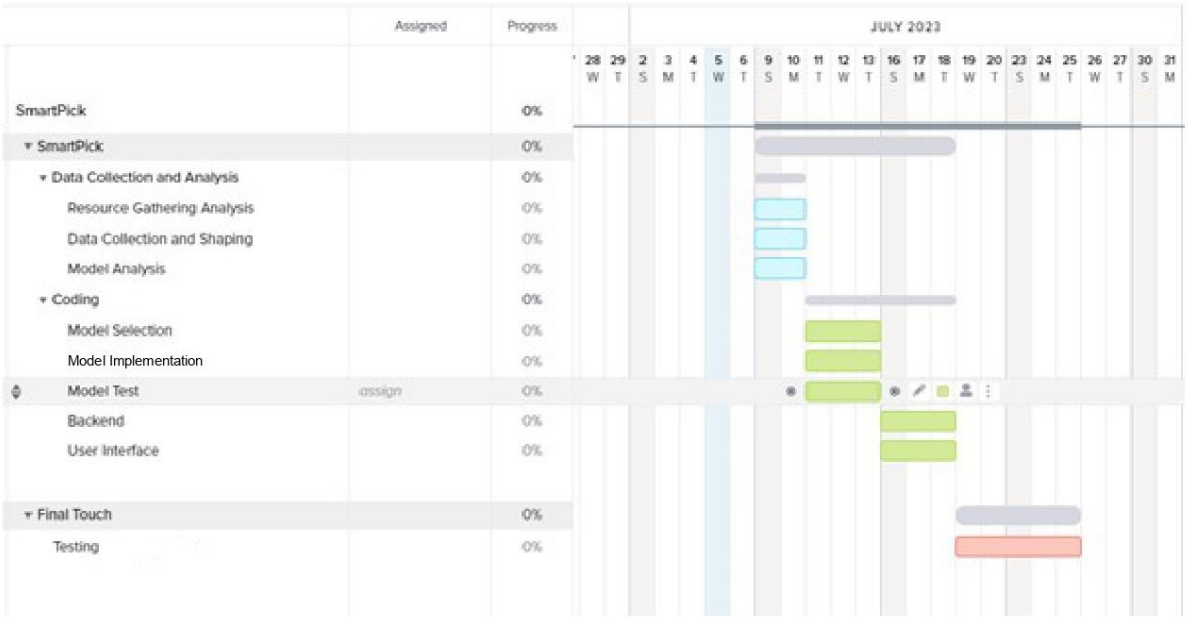


Figure 1: Timeline of the project

3.4 Process/Activity wise Resource Allocation

Resources, including team members' time and computing resources, were allocated based on the complexity of tasks and individual expertise. The team efficiently managed the provided resources by utilizing open-source tools and libraries for ML model training and data engineering.

3.5 Estimated Costing

The internship program, funded by the Bangladesh government through the EDGE program, provided the necessary financial support for the project's execution. The financing covered expenses related to hardware, software and other essential components. To ensure cost-effectiveness, the team explored free or low-cost options for data storage and project management tools. This approach enabled the team to focus on the project's core development without incurring unnecessary expenses.

Chapter 4

Methodology

The development of the AI-based Book Recommendation System with popularity-based and collaborative filtering techniques, integrated with Django as the backend framework, involved several key steps. The methodology can be outlined as follows:

- Data collection and Preprocessing: The first phase involves data collection, where a comprehensive dataset containing information about books, such as title, author, genre, and ratings, along with user data containing reading history, was obtained. The dataset was then preprocessed to handle missing values, normalize ratings, and transform categorical features into a suitable format for the AI model.
- AI Model: The AI model was trained on the preprocessed dataset to learn users' preferences and identify patterns among books. The models incorporated the following:

    - Popularity-Based Recommendation Technique: This involved analyzing the frequency of ratings on each book by the users and averaging them. Then selecting the top-rated books as popular recommendations. The popularity scores were computed and stored in the local storage.
    - Collaborative Filtering with Cosine Similarity: The algorithm computes the cosine similarity of all the books in the database with all other books. In return, for each book, the system suggests other books that share substantial similarities. This collaborative filtering approach ensures that users receive recommendations that align closely with their specific interests.

- Django Backend Integration: Django was used as the backend framework to store the preprocessed data in a database and handle user interactions. The AI model was integrated into the Django backend to fetch recommendations based on user requests and provide real-time personalized book suggestions. Django's efficient data handling capabilities facilitated smooth communication between the AI model and the front-end interface.

- Frontend Development: The frontend of the Book Recommendation System was developed using HTML, CSS, and JavaScript, providing an intuitive and user-friendly interface. The front end allowed users to interact with the system, view popular book recommendations, and receive personalized suggestions based on their preferences. The frontend communicated with the Django backend to fetch and display the relevant data.
- Evaluation and Testing: The evaluation of the system involved thorough testing and feedback collection from users.

Chapter 5

Body of the Project

5.1 Work Description

Throughout my internship at ADN Diginet, I had the opportunity to engage in an innovative approach to training and project advancement. ADN Diginet, understanding the profound impact of AI and ML, had introduced an extensive training initiative. The aim was to furnish interns with the expertise and proficiencies required to participate in ML-driven solutions that resonate with the company's forward-thinking perspective. Rather than engaging in immediate production tasks, we were fully immersed in a demanding training program. Our goal was to develop a demonstrative project that prominently demonstrated the capabilities of AI.

We began by building a system from zero for recommending books. The prototype system was developed to show the potential of combining a popularity-based recommendation system with cosine similarity-based collaborative filtering. The project's primary focus was machine learning, but it also had a strong e-commerce aspect. By putting orders, evaluating books, as well as adding books to their carts, users could interact with the system. However, the payment method was absent. My responsibility involved developing the ML model for the project, which included feature engineering, popularity-based recommendation, data preparation, and the use of the cosine-similarity method to suggest related books.

5.2 Requirement Analysis

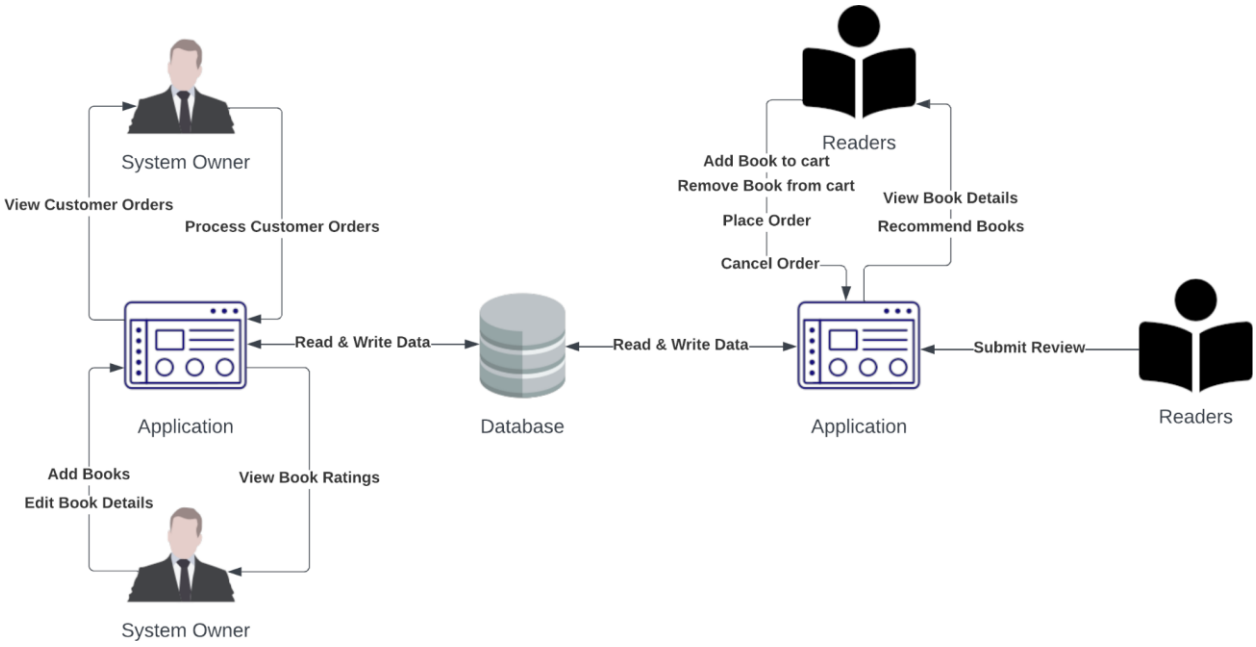


Figure 2: Rich Picture of the system

Functional Requirements

- **Add new book or books (System Owners):**
    - The system owner needs to be able to add new book or books to the collection using the system.
    - All the necessary information about the book's such as title, authors, ISBN etc. are all required fields.

- **Alter details of books (System Owners):**
    - The ability to change current book details should be given to system owners.
    - The system needs to verify and update the latest information while ensuring data integrity.

- **View book details:**
    - Each book should have detailed information available to readers and also the suggestion for similar books.

- **Delete books from own cart (Readers):**
    - It should be possible for readers to take books out of their carts.

- **View ratings of a user to a book (System Owners):**

    - Every user's evaluation of every book in the system's collection needs to be visible to the system owners.
    - It is important to present ratings in an intuitive and organized manner.

- **Recommend books:**

    - Based on the ratings and popularity index of their previously purchased books, readers should receive individualized book recommendations.
    - The reader's interests should be considered while making a recommendation.

- **Add book or books to own cart:**

    - It should be possible for readers to add books to their shopping basket for later purchase.

Non-Functional Requirements

- **Security:**

    - Information about users should be protected and data security should be guaranteed.
    - Mechanisms for user authentication and authorization need to be in place.

- **Performance:**

    - The system must respond quickly and support a manageable number of concurrent users.
    - Algorithms for making recommendations should deliver findings quickly and accurately.

- **Scalability:**

    - If the quantity of books and users increases, the system should be scalable.
    - It should be able to handle an expanding amount of data.

- **Usability:**

    - Both system owners and readers should find the user interface to be simple and easy to use.

- It is important to provide clear labeling and proper navigation.

- **Data Accuracy:**

  - Data loss or corruption should be prevented by the system by safeguarding data security.
  - Implementing proper data validation and error management is necessary.

- **Compatibility:**

  - The system needs to support the common web browsers and devices that readers and system owners use.
  - It is important to perform cross-browser and cross-device testing.

- **Response time:**

  - When a user adds something to their shopping cart or submits a rating, the system should react immediately.
  - Response times must be within reasonable bounds.

- **Privacy:**

  - The system must uphold data protection laws and preserve user privacy.
  - Privacy rules should be followed, and user data should be handled securely.

5.3 System Analysis

5.3.1 Six Element Analysis

Table 1: Six Element Analysis of the system

Process	System Roles					
	Human	Non-Computing Hardware	Computing Hardware	Software	Database	Networking & Communication
Add Books	**System Owner:** System owners will be able to add a new book to the database	**Book:** The book whose data will be added. **Paper:** Book details might come	**Computer:** Computer or smartphone might be used to access the system or	**Book Recommendation & ECommerce System:** To input new book details.	**MySQL:** Store book details, user details, user orders, and buying history, user	**Internet:** To access the *book recommendation and ecommerce system*.

	through the django admin panel.	printed on a sheet of paper.	read or create book details in word processing software or spreadsheet software that will be added to the system. **Printer:** It might be used to print out book details.	**Office Suite:** Read or write word document or spreadsheet containing book details.	reviews.	
Edit Book Details	**System Owner:** System owners will be able to edit book details in the database through the django admin panel.	**Book:** The book whose data will be edited. **Paper:** Book details might come printed on a sheet of paper.	**Computer:** Computer or smartphone might be used to access the system or read or create book details in word processing software or spreadsheet software that will be edited in the system. **Printer:** It might be used to print out book details.	**Book Recommendation & ECommerce System:** To edit book details. **Office Suite:** Read or write word document or spreadsheet containing book details.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system*.
View Book Ratings	**System Owners:** System Owners will be able to view individual reader's	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone	**Book Recommendation & ECommerce System:** To view book ratings.	**MySQL:** Store book details, user details, user orders, and buying history, user	**Internet:** To access the *book recommendation and ecommerce system*.

	ratings to book through the django admin panel. **Readers:** Readers will be able to view the average rating of a book in the book details page.		to access the system.		reviews.	
View Customer Orders	**System Owners:** System Owners will be able to view individual reader's orders through the django admin panel.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To view customer orders.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system*.
Process Customer Orders	**System Owners:** System owners will be able to process orders and initiate different processes for the delivery system, like fetching books from the warehouse, handing over to delivery service and updating the order as delivered upon delivery was completed	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To process customer orders.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system*.

	successfully.					
Recommend Books	**Readers:** Readers will be able to view book recommendations on their homepage and individual book pages.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To recommend books to the readers.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system*.
View Book Details	**Readers:** Readers will be able to view book details like book title, authors, book summary, average book rating, book price, adding book to the cart and related books generated by the book recommendation system.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To view book details by the readers.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system*.
Add Books to Cart	**Readers:** Readers will be able to add books to the	Not Applicable	**Computer:** System owners and readers will	**Book Recommendation & ECommerc**	**MySQL:** Store book details, user details, user	**Internet:** To access the *book recommendati*

	cart from the homepage and/or book details page.		need a computer or smartphone to access the system.	**e System:** To add books to the cart.	orders, and buying history, user reviews.	*on and ecommerce system.*
Remove Books from Cart	**Readers:** Readers will be able to remove books from the cart from the cart page.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To remove books to the cart.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system.*
Place Order	**Readers:** Readers will be able to place an order of the books that have been added to the cart from the cart from the cart page.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To place book orders.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system.*
Cancel Order	**Readers:** Readers will be able to cancel the order that has not been placed yet from the cart from the cart page.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To cancel an order.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system.*
Submit Review	**Readers:** Readers will be able to rate books that have been delivered from the cart page.	Not Applicable	**Computer:** System owners and readers will need a computer or smartphone to access the system.	**Book Recommendation & ECommerce System:** To submit book ratings.	**MySQL:** Store book details, user details, user orders, and buying history, user reviews.	**Internet:** To access the *book recommendation and ecommerce system.*

5.3.2 Feasibility Analysis

The project's technical feasibility is essential, mostly because of the strategic technology choices made. The smooth integration of ML capabilities into the project is made possible by Python's broad support for ML frameworks like scikit-learn, TensorFlow, and PyTorch. Furthermore, given Python's significance in the ML field and the availability of numerous libraries for ML integration, it was an obvious choice for the backend. The technical feasibility is improved by Django's web framework thanks to its scalability, security features, and developer-friendliness. Particularly when managing e-commerce functionality and user data, Django's built-in stability and security features deliver an extra level of confidence. This guarantees that the technical base is both strong and capable of satisfying performance and security needs.

5.3.3 Problem Solution Analysis

The project aimed to suggest books that would be a good fit for each user. Traditional recommendation systems sometimes didn't quite get it right, so we brought in machine learning models, especially one called collaborative filtering using cosine similarity. This helped us understand what users like and how they behave, so we could give them better recommendations.

Choosing Python and Django for the technical part was important. It helped us build a strong foundation. This project wasn't just about solving the book suggestion problem. It became a great example of how powerful machine learning can be in the real world. It got our team excited about learning and sharing knowledge.

This project showed us how machine learning can make a big difference in all sorts of real-world challenges. It's not just about suggesting books—it's about making user experiences better, matching up with what the organization wants, and pushing the boundaries of innovation in many different areas.

5.3.4 Effect and Constraints Analysis

The project was essentially an implementation primarily designed to serve as a demonstrator of ML capabilities. It was initiated from the ground up, with the advantage of not having to deal with any existing components that could be disrupted during implementation. However, while the project's core functionality was stable, questions might arise regarding its adaptability for future modifications or the incorporation of additional features. This is because the primary objective was to speed up the creation of a demo within the constraints of a limited timeframe.

Regarding constraints, our project faced limitations in both time and data collection. We had a strict one-month timeline for implementation, necessitating a rapid development process. Furthermore, because the project was essentially a demonstration, an ample or adequate data set was not available to us.

5.4 System Design

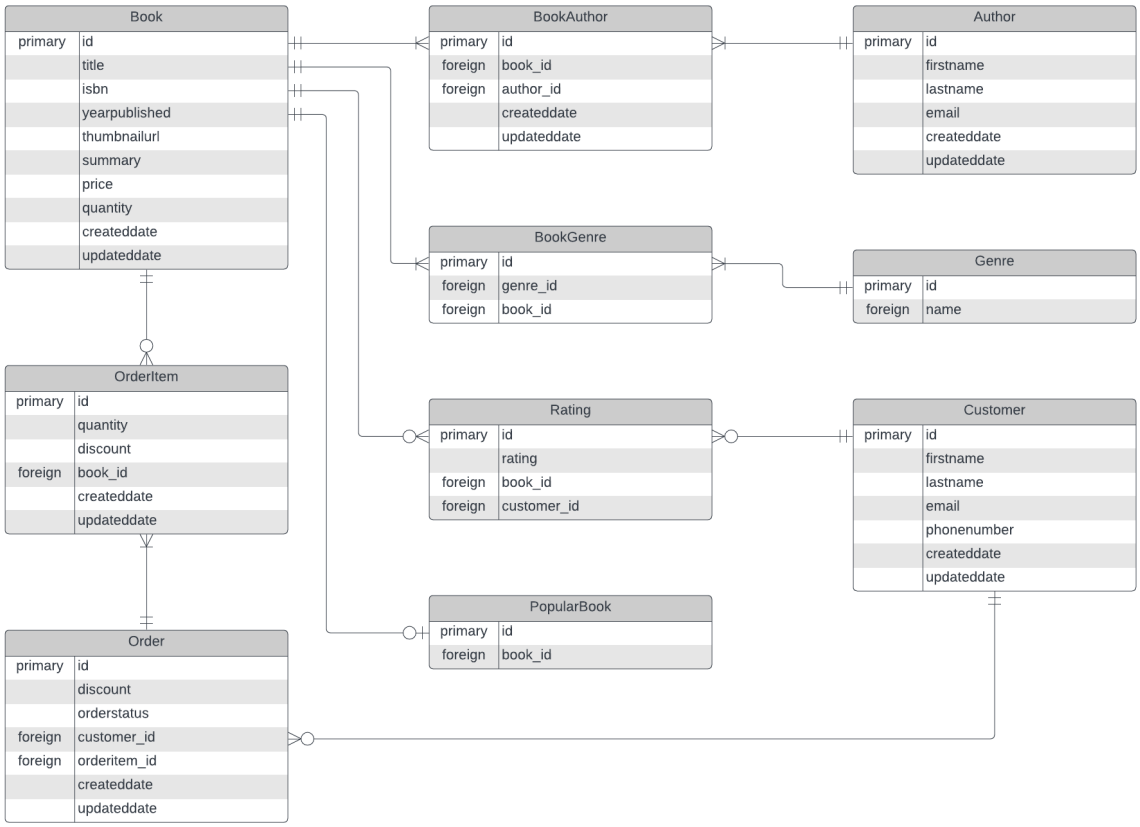


Figure 3: Entity Relationship Diagram (ERD) of the system

5.5 Implementation

We used the book recommendation dataset from Kaggle which consisted of 3 csv files – books.csv, users.csv and ratings.csv. Then into jumped into the phase of data preprocessing and model creation.

Popularity based recommender model:

In this model, we considered a simple approach by finding the top 50 books with the highest average rating but we choose only those books which have got a minimum of 250 votes so that system outputs meaningful data. We merged the ratings dataset with books dataset on ISBN number of the books and used group by on book title to count the number of votes on each book to create a new data frame. We did not use ISBN for group by because a book can have multiple ISBN numbers hence we considered the book title. Then we created another data frame to find the average ratings of each book, adding all the ratings divided by the total number of ratings. Then we merged the two data frames which consisted of the book titles, number of votes on the books and the average rating of the books. Afterwards, we filtered out only those books which had a total vote of minimum 250. We sorted the data frame in descending order on top on average ratings to get the most popular books on the top and the least books on the bottom. Then we considered the first 50 books to display to the user. Next, we needed the books details of the books in our resultant data frame. So, we merged our data frame with books dataset on book title column and dropped duplicates since a book title can have multiple ISBN numbers. Finally, we had the data of the most popular books to display to the user.

Collaborative-Based filtering model:

In this technique, we used the cosine-similarity on the books data frame to find the Euclidean distance of each book with all other books. The Euclidean distance determines how similar a book is to another book. Before using the cosine-similarly we needed to create a books data frame where each book could be represented by a vector. To create the book data frame, first we created a data frame where in the columns all the users were present and in the index we had our books and in the grid we had the data of ratings of each book by each of the user. To create an intelligent system, we considered two criteria on the data frame. We choose only

those users who have rated a minimum of 200 books and those books which had minimum of 50 votes. Finally, we had the data frame of books where each book was represented by a vector. Then we used the cosine-similarity on the data frame to find the Euclidean distance or the similarity score of each book with all other books. The lower the Euclidean distance the greater the similarity score hence the more similar the books. Lastly, we created a function which returns the top 5 (the value can be changed) similar book to the input book name.

Then to display the recommendations to the users, we exported the popular books data frame and the similarity scores data frame as pickle files. The files were used by the backend team to send the requested data to frontend.

5.6 Testing

Input and Output

```
popular_books
[47]  ✓ 0.0s                                                    Python
...
{0: {'Book-Title': 'Harry Potter and the Prisoner of Azkaban (Book 3)',
  'Book-Author': 'J. K. Rowling',
  'Image-URL-M': 'http://images.amazon.com/images/P/0439136350.01.MZZZZZZZ.jpg',
  'num_ratings': 428,
  'avg_rating': 5.85},
 3: {'Book-Title': 'Harry Potter and the Goblet of Fire (Book 4)',
  'Book-Author': 'J. K. Rowling',
  'Image-URL-M': 'http://images.amazon.com/images/P/0439139597.01.MZZZZZZZ.jpg',
  'num_ratings': 387,
  'avg_rating': 5.82},
 5: {'Book-Title': "Harry Potter and the Sorcerer's Stone (Book 1)",
  'Book-Author': 'J. K. Rowling',
  'Image-URL-M': 'http://images.amazon.com/images/P/0590353403.01.MZZZZZZZ.jpg',
  'num_ratings': 278,
  'avg_rating': 5.74},
 9: {'Book-Title': 'Harry Potter and the Order of the Phoenix (Book 5)',
  'Book-Author': 'J. K. Rowling',
  'Image-URL-M': 'http://images.amazon.com/images/P/043935806X.01.MZZZZZZZ.jpg',
  'num_ratings': 347,
  'avg_rating': 5.5},
```

Figure 4: Data of top 50 popular books (Image is cropped)

The screenshot above shows the top 50 books (four are incorporated in the screenshot) generated by the popularity-based recommendation model. The data frame contains the book-id as index and the columns represent the book-title, image, number of ratings and average

ratings of each book. The average rating of the first book is the highest and it decreases as we go down.

```
recommend('1984')
```

[49] ✓ 0.4s

```
[['Animal Farm',
  'George Orwell',
  'http://images.amazon.com/images/P/0451526341.01.MZZZZZZZ.jpg'],
 ["The Handmaid's Tale",
  'Margaret Atwood',
  'http://images.amazon.com/images/P/0449212602.01.MZZZZZZZ.jpg'],
 ['Brave New World',
  'Aldous Huxley',
  'http://images.amazon.com/images/P/0060809833.01.MZZZZZZZ.jpg'],
 ['The Vampire Lestat (Vampire Chronicles, Book II)',
  'ANNE RICE',
  'http://images.amazon.com/images/P/0345313860.01.MZZZZZZZ.jpg'],
 ['The Hours : A Novel',
  'Michael Cunningham',
  'http://images.amazon.com/images/P/0312243022.01.MZZZZZZZ.jpg']]
```

Figure 5: Output of recommend function when "1984" is send as parameter

The above screenshot shows the out of a function name “recommend” which we talked about in the previous section. The function outputs 5 similar books when we pass a book name as parameter.

```
recommend('A Bend in the Road')
```

[50] ✓ 0.2s

```
[['A Walk to Remember',
  'Nicholas Sparks',
  'http://images.amazon.com/images/P/0446608955.01.MZZZZZZZ.jpg'],
 ['The Last Time They Met : A Novel',
  'Anita Shreve',
  'http://images.amazon.com/images/P/0316781266.01.MZZZZZZZ.jpg'],
 ['Sea Glass: A Novel',
  'Anita Shreve',
  'http://images.amazon.com/images/P/0316089699.01.MZZZZZZZ.jpg'],
 ['Angels',
  'Marian Keyes',
  'http://images.amazon.com/images/P/0060008024.01.MZZZZZZZ.jpg'],
 ['Family Album',
  'Danielle Steel',
  'http://images.amazon.com/images/P/0440124344.01.MZZZZZZZ.jpg']]
```

Figure 6: Output of recommend function when “A Bend in the Road” is send as parameter

Chapter 6

Results & Analysis

The project demonstrated a successful application of machine learning (ML) in a practical setting, particularly in crafting a personalized book recommendation system. One significant accomplishment was that it served as a powerful example of machine learning's potential, as demonstrated by the skillful application of ML methods such as collaborative filtering with cosine similarity. Through personalized book recommendations, this not only transformed user engagement but also opened the door for creative thinking in a variety of industries. The system excelled in recommending timeless classics to users and providing similar or related book suggestions based on user searches. Additionally, the project substantially augmented the skill set of team members, offering them hands-on proficiency in integrating ML models, utilizing Python and Django, and fostering cross-disciplinary collaboration.

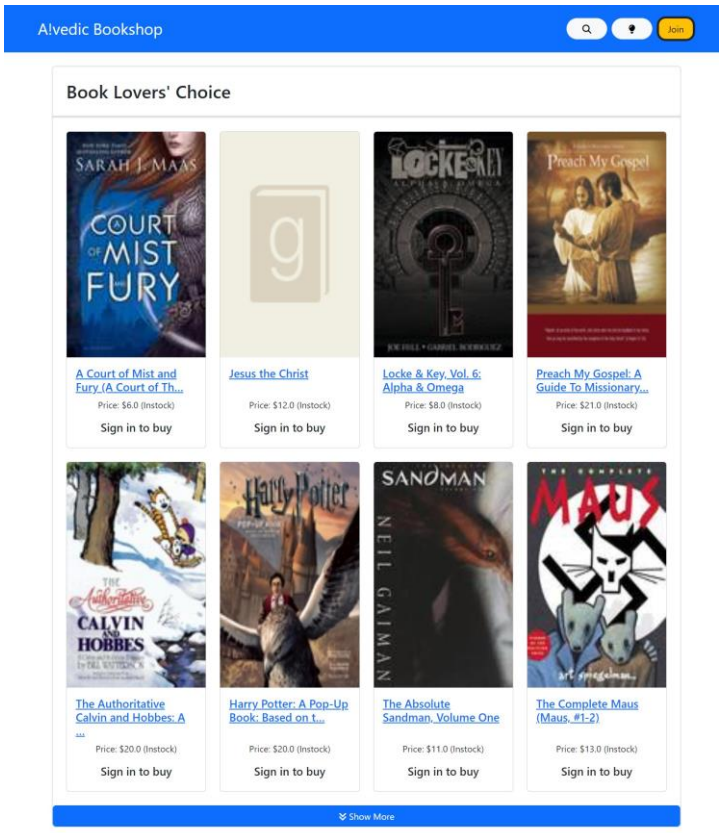


Figure 7: Displaying the popular books in UI

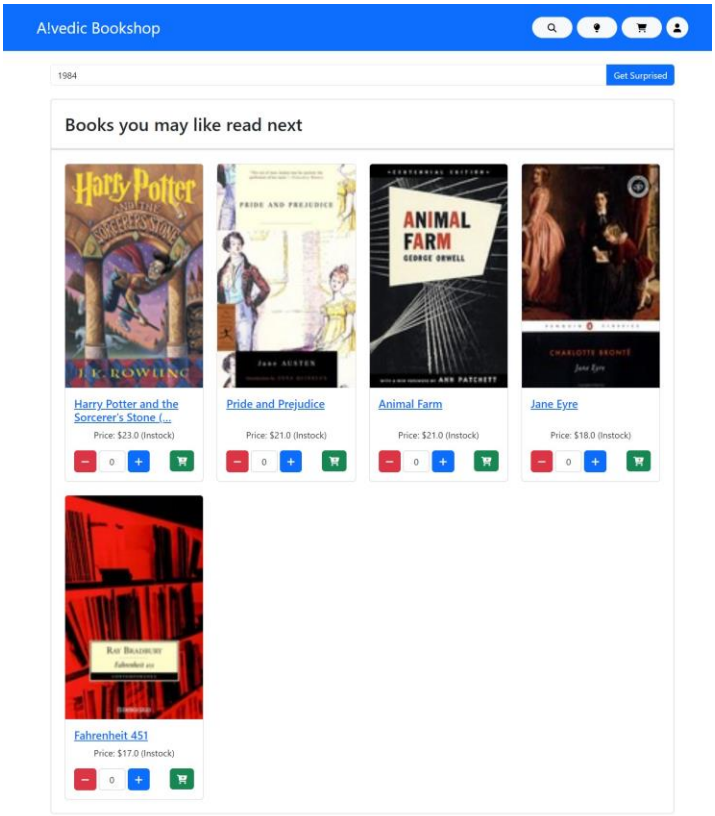


Figure 8: Displaying the similar books when a book name is given as input in UI

In addition, the project was successfully completed within the constrained one-month timeline, demonstrating effective resource allocation and project management. This accomplishment demonstrated the team's capacity to work under pressure and produce a workable machine learning solution. However, given the present project would encounter difficulties if it were to grow to support greater user populations or larger datasets, future initiatives ought to assess scalability requirements.

One area for consideration is the project's alignment with ADN Diginet's product scope. Designed primarily as a demo, it did not directly integrate with the company's existing product offerings. While this approach allowed for innovative exploration, future projects should consider how ML applications can directly enhance ADN Diginet's product line, ensuring a more seamless integration of ML capabilities. In summary, the project achieved its primary objective of showcasing ML's potential in a real-world context, delivering valuable insights, enhanced skills, and a culture of knowledge sharing within ADN Diginet.

Chapter 7

Project as Engineering Problem Analysis

7.1 Sustainability of the Project/Work

The project didn't just bring immediate benefits but it left a long-lasting mark for future systems. It helped our team learn a lot about using advanced technology like machine learning. This new knowledge will stick around and be useful for future projects. The project also encouraged us to work together with different parts of the company and share what they know. It showed how we can use ML to solve real-world problems, and this is something we can keep doing in the future. The ideas from this project will help ADN DigiNet plan and do even bigger projects with machine learning.

7.2 Social and Environmental Effects and Analysis

The project aimed to utilize machine learning for better user experiences, leading to significant social and environmental effects. It notably enhanced user satisfaction and engagement within the book readers community. Additionally, it fostered knowledge sharing and teamwork among team members, promoting a collaborative work culture. On the environmental front, the project's emphasis on digital solutions reduced the need for physical resources, aligning with eco-friendly practices and minimizing environmental impact.

7.3 Addressing Ethics and Ethical Issues

The project, all about making users' experiences better with machine learning, brings up some important ethical issues. Keeping data safe and getting user permission is really important. We also need to be careful about any unfairness in machine learning, especially in suggesting

things, so that everyone gets treated the same. Using AI in an ethical way means being clear about what we're doing, taking responsibility, and making sure users understand how it works. These rules aren't just for this project, but for using machine learning in general. Finally, it's important to keep checking and making sure we're following the right ethical guidelines, because they might change over time.

Chapter 8

Lesson Learned

8.1 Challenges Faced During this Period

Gathering Data

Building a large, high-quality dataset for machine learning can be challenging. Obtaining sufficient and diverse types of data, verifying the accuracy of the data, and managing noisy or incomplete data were some of the problems. To build a web application that utilizes the machine learning model and suitable user interface, we had to comb through a substantial collection of real-world datasets that held the data we needed for our use case.

Data Preprocessing and Filtering

The dataset preparation process for training may take some time. Errors may arise from improper handling of data cleaning and preprocessing tasks such as addressing outliers, inserting missing values, and scaling features.

Challenges with Feature Engineering

Establishing features that are relevant and useful is necessary for ML models to function successfully. Determining which characteristics are most crucial and how to design them to capture critical data is the challenging part.

Model Selection Dilemmas

Choosing the most suitable ML model can be tricky. Different algorithms have different strengths and weaknesses, and selecting the wrong one may lead to suboptimal results. Proper model selection involves thorough evaluation and comparison. We also had the time constraint, so we had to choose a tried and tested model that had been utilized in the relevant fields and one that code was implemented in a short amount of time and requires relatively less powerful model training workstations. Therefore, our choice was to use a cosine similarity-based recommendation system.

Hyperparameter Tuning

During the model training phase, we exercised meticulous discretion in selecting the data from the dataset. We deliberated on which reader ratings should be prioritized over others and which books should be included in the training process to ensure that the resulting recommendations from our trained models are coherent and meaningful. Additionally, we took care to prevent overfitting or underfitting of the model, aiming for strong generalization capabilities across the dataset. This was done with the overarching goal of enhancing the predictive accuracy of the model.

Chapter 9

Future Work & Conclusion

9.1 Future Works

In the context of this work, there are a lot of fascinating opportunities for future work. One primary focus area is improving recommendation algorithms. This involves using other recommendation techniques like content-based recommendations and hybrid models. Additionally, the project can evolve towards providing dynamic and real-time recommendations. A chat-bot would increase the user interaction with the system.

It is nevertheless a necessity that ethical issues be continuously considered. In order to guarantee that recommendations are unbiased, further work might include the application of fairness-aware algorithms. Furthermore, gathering and evaluating user input to evaluate the caliber of recommendations and user happiness is a continuous procedure. As data sizes increase, scalability and speed optimization become critical. In order to maintain the effectiveness and responsiveness of recommendations, future study may involve investigating distributed computing options and optimizing algorithms for efficiency.

9.2 Conclusion

ADN's internship program in artificial intelligence (AI), backed by Bangladesh's EDGE program, has been valuable. We learned a lot about ML, like different types and models. We started with classes and hands-on activities led by experts. We also worked together to find real-world problems that ML could solve, which was a big highlight. We made a working version of a book recommendation system discussed in this entire report. It showed how good planning and teamwork can make a project succeed. It proved that ML can be super useful in real life. We also learned how important it is to manage a project well. We picked up skills like cleaning up data, choosing the right models, and making sure everything was fair and ethical. Looking ahead, we can make the recommendation system even better and faster. We're committed to learning more and making sure ML is used in a good way. Overall, this internship taught us a ton about how ML works in the real world. It's a big step towards cool careers in ML and AI, and it's helping technology grow in Bangladesh and beyond.

Bibliography

[1] Da'u, A., Salim, N. Recommendation system based on deep learning methods: a systematic review and new directions. *Artif Intell Rev* 53, 2709–2748 (2020).

[2] Emanet, Nahit & Öz, Halil & Bayram, Nazan & Delen, Dursun. (2014). A comparative analysis of machine learning methods for classification type decision problems in healthcare. Decision Analytics. 1. 6. 10.1186/2193-8636-1-6.

[3] Karimi, Zohreh. (2021). Popularity Based Filtering Recommender Systems.

An Undergraduate Internship/Project on Book Recommendation System at ADN DigiNet

By

Ashraful Kabir

Student ID: **2022700**

Summer, 2023

Consent from Supervisor

The student modified the internship final report as per the recommendations made by his/her academic supervisor and/or panel members during and/or before final viva, and the department can use this version for archiving as well as the OBE course material for CSE499.

This internship report is checked with Turnitin or Ithenticate plagiarism checker, and the score is:

Turnitin Score (%): 14

.....Marzan.....

Marzan Binte Hasan

Department of Computer Science & Engineering

Independent University, Bangladesh