



Green University of Bangladesh

*Department of Computer Science and Engineering (CSE)
Semester: (Fall, Year: 2023-24), B.Sc. in CSE (Eve)*

Project Name: Content Based Movie Recommendation System

*Course Title: Artificial Intelligence Lab
Course Code: CSE-316
Section: 2012-EB*

Students Details

Name	ID
Md. Mahmudul Hasan	201015083
S.M. Mustafizur Rahman	212015052

*Submission Date: 05/01/2024
Course Teacher's Name: Tanoy Debnath*

[For teachers use only: **Don't write anything inside this box**]

<u>Lab Project Status</u>	
Marks:	Signature:
Comments:	Date:

Contents

1	Introduction	3
1.1	Overview	3
1.2	Related works	4
1.3	Problem Domain and Motivation	4
1.4	Problem Statement	4
1.5	Design Goals/Objectives	5
1.6	Project Planning	5
1.7	Project Features	5
2	Design/Development/Implementation of the Project	6
2.1	Introduction	6
2.2	Project Details	6
2.2.1	Scope of the Project	6
2.2.2	Literature Survey	7
2.2.3	Methodology for Movie Recommendation	7
2.2.4	System Architecture Design	8
2.3	Implementation	8
2.3.1	Movie Recommendation System	8
2.3.3	Tools and Technologies	8
2.3.2	Favourite Movie Search	9
2.3.4	Movie Recommendation System Code	10
2.4	Gantt Chart	10
3	Performance Evaluation	11
3.1	Results Analysis/Testing	11
3.1.1	Unit Testing	11
3.1.2	System Testing	11

3.2	Result and performance Analysis	11
3.3	Comparison(Table)	12
4	Conclusion	13
4.1	Discussion	13
4.2	Limitations	13
4.3	Scope of Future Work	14

Chapter 1

Introduction

1.1 Overview

- A content-based movie recommendation system suggests items to users based on the features of the items and preferences expressed by the user.
- In the context of movies, this means recommending movies to users based on the movie's content or features themselves, rather than relying on other users' preferences.
- Here's an overview of how a content-based movie recommendation system typically works.
- **Data Collection** Collect information about movies. Collect data on user preferences if available.
- **Feature extraction** Extract relevant features from movie data. If user preferences are available, extract relevant attributes.
- **Representing movies and users as vectors** in vectorized feature space.
- **Similarity** is calculated by measuring the similarity between items based on their feature vectors.
- **Ranking** ranks items based on their similarity score. Recommends top items to the user
- **Deploy** the recommendation system by integrating it with the user interface of the implemented streaming platform or application.
- In practice, hybrid systems that combine content-based and collaborative filtering approaches are often used to take advantage of the strengths of both approaches.

1.2 Related works

References	Methodology	Findings	Dataset
K. Sparck Jones	Water Fall	Feature Importance	MovieLens Dataset
Pradeep NK	Neural Network Model	Text Data Processing	IMDb Datasets
G. Salton	CNN Model	Limited Diversity	Kaggle Datasets

1.3 Problem Domain and Motivation

- Our project entitled “Content Based Movie Recommendation System” aims to suggest or recommend the various users, the movie they might like, by intake of their ratings, genres and similarities.
- The system proposed is a kind of content based filtering system which finally recommends the likable movie to the users using cosine similarity algorithm.
- This will extract vital information and recommend the users according to user’s preferences, interest, or history about movies.
- Our system is to use dataset which are to be thoroughly filtered in order to gain user’s idea for movies.
- This system is to be implemented with machine learning by python. This filtering method matches content resources to user characteristics, base their predictions on user’s information.
- It relies heavily on the ratings of different users.
- Purpose of movie recommendation system aims to provide users with accurate movie recommendations.
- Usually basic recommendation system to make recommendations consider one of the following factors User preference known as content based Filtering or the preference of similar users known as collaborative filtering.

1.4 Problem Statement

- The goal of the project is to recommend a movie to the user.
- Providing related content out of relevant and irrelevant collection of items to users of online service providers

1.5 Design Goals/Objectives

- Recommender systems are information filtering tools that aspire to predict the rating for users and items, predominantly from big data to recommend their likes.
- Movie recommendation systems provide a mechanism to assist users in classifying users with similar interests. The purpose of a recommendation system basically is to search for content that would be interesting to an individual.
- Moreover, it involves a number of factors to create personalized lists of useful and interesting content specific to each user individual.
- The Objectives of the project is: -
 - Recommend movie to the user Saving time of the user.
 - Improving the Accuracy of the recommendation system.
 - Improve the Quality of the movie Recommendation system.
 - Improving the Scalability.
 - Enhancing the user experience.
 - A report summarizing exploratory analysis results, recommendation algorithms used and predicted user-movie rating
 - Influencing usage pattern in users.

1.6 Project Planning

- 1.Data Collection 2.Programming Language 3.Data Preprocessing 4.Feature Extraction 5.User Profile Creation 6.Similarity Calculation 7.Evaluation 8.Model Improvement 9.Deployment.

1.7 Project Features

- 1.User Profile Creation 2.Movie Feature Extraction 3.Similarity Calculation 4.Search Functionality 5.User History Tracking 6.combined features.

Chapter 2

Design/Development/Implementation of the Project

2.1 Introduction

- A recommendation system or recommendation engine is a model used for data files where it tries to predict a user's preferences. These systems have become increasingly popular these days and are widely used in movies, music, books, videos, clothing, restaurants, food, places. These systems collect information about user preferences and behavior, and then use this information to improve their recommendations in the future. There are different types of movies such as some for entertainment, some educational, some for children, animated movies, and some horror movies or action movies.
- Movies can be easily separated by comedy, thriller, etc. Animation, Action etc. Our most popular movies movie recommendation system helps us search for our favorite movies with a variety of movies among these movies and therefore reduces the hassle of spending many times searching for our favorite movies. So, it is necessary that movie recommendation system should be very reliable. A large number of companies are using recommendation system to increase user interaction and enrich users. Recommendation systems have several benefits, the most important being customer satisfaction and revenue. Movie recommendation system is very powerful and important system

2.2 Project Details

2.2.1 Scope of the Project

- The objective of this project is to provide accurate movie recommendations to users. The goal of the project is to improve the quality of movie recommendation system, such as accuracy, quality and scalability of system than the pure approaches. This is done using content based filtering , To eradicate the overload of the data, recommendation system is used as information filtering tool in social networking sites. Hence, there is a huge scope of exploration in this field for improving scalability, accuracy and quality of movie recommendation systems Movie Recommendation system is very powerful and important system. But, due to the

problems associated with pure collaborative approach, movie recommendation systems also suffers with poor recommendation quality and scalability issues.

2.2.2 Literature Survey

- After the study of recommending items from some fixed database has been done, two main recommending techniques have emerged which are content-based technique and collaborative technique.
- In content-based recommendation, items are recommended which are similar to those provided by the user, whereas in collaborative recommendation users whose tastes are similar are identified to those of the given user and recommends items they have liked.
- Later with the evolution of the recommender system hybrid method has been invented which merges two or more techniques.
- Before the invention of the recommending system, people had to read reviews and choose the movie that best suited their interest or had to randomly choose any movie based on some other criteria.
- This became difficult as the number of movies that are available online started increasing rapidly.

2.2.3 Methodology for Movie Recommendation

- The proposed movie recommendation system gives finer similarity metrics and quality than the existing Movie recommendation system. The proposed approach is for improving the scalability and quality of the movie recommendation system. We use a Hybrid approach, by unifying Content-Based Filtering. For computing similarity between the different movies in the given dataset efficiently and in least time and to reduce computation time of the movie recommender engine we used cosine similarity measure.
- Water Fall Methodology
- Water Fall Development methodology is one of the best software development approaches.
- Collecting the data sets: Collecting all the required data set from Kaggle web site. In this project we require movie.csv.
- Data Analysis: make sure that the collected data sets are correct and analysing the data in the csv files. i.e. checking whether all the column fields are present in the data sets.
- Algorithms: in our project we have only one algorithm one is cosine similarity used to build the machine learning recommendation model.

- Training and Testing the model: Once the implementation of algorithm is completed. we have to train the model to get the result. We have tested it several times the model is recommend different set of movies to different users.
- Improvements in the project: In the later stage we can implement different algorithms and methods for better recommendation.

2.2.4 System Architecture Design

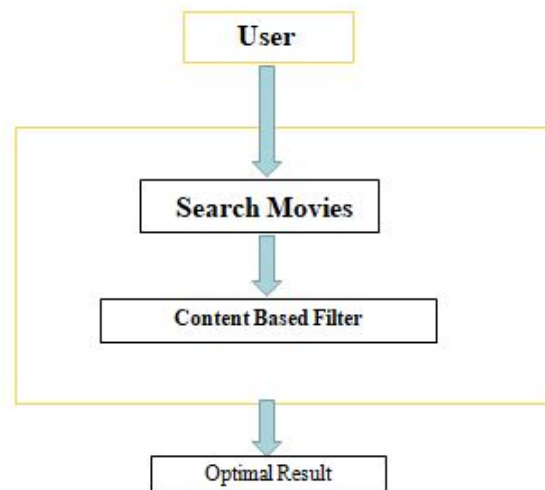


Figure 2.1: System Architecture

2.3 Implementation

- Implementing a content-based movie recommendation system involves several steps. We will implement Sample Movie Data, User Profile, Feature Extraction and Vectorization, Similarity, Display Recommended Movies to use Python and Skit-Learn library for vectorization and similarity calculation.

2.3.1 Movie Recommendation System

2.3.3 Tools and Technologies

- Building a Content-Based Movie Recommendation System involves several tools and technologies. Here's a list of commonly used tools and technologies for different stages of the recommendation system development The tools and technologies we use in this project is:-
- 1.Google Colab. 2.Machine learning. 3.Cosine similarity Algorithm. 4.Python Code. 5.Dataset. 6.Numpy. 7.Pandans. 8.Difflib. 9.Sklearn. 10.Seaborn. 11.Matplotlib.

2.3.2 Favourite Movie Search



Figure 2.2: Movie Search

2.3.4 Movie Recommendation System Code

```

movie_name = input(' Enter your favourite movie name : ')

list_of_all_titles = movies_data['title'].tolist()

find_close_match = difflib.get_close_matches(movie_name, list_of_all_titles)

close_match = find_close_match[0]

index_of_the_movie = movies_data[movies_data.title == close_match]['index'].values[0]

similarity_score = list(enumerate(similarity[index_of_the_movie]))

sorted_similar_movies = sorted(similarity_score, key = lambda x:x[1], reverse = True)

print('Movies suggested for you : \n')

i = 1

for movie in sorted_similar_movies:
    index = movie[0]
    title_from_index = movies_data[movies_data.index==index]['title'].values[0]
    if (i<30):
        print(i, '.',title_from_index)
        i+=1

```

Figure 2.3: Python Code

<https://drive.google.com/file/d/1qNU0jdBKzlc5xaGkpk9J-hFhV-7oVtF/view?usp=sharing>

2.4 Gantt Chart

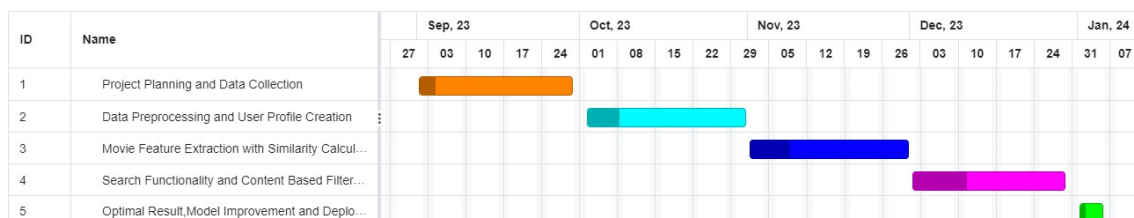


Figure 2.4: Gantt Chart

Chapter 3

Performance Evaluation

3.1 Results Analysis/Testing

- There are many different types of testing methods. Some of the important testing methodologies are:

3.1.1 Unit Testing

- Unit testing is the first level of testing and is often performed by the developers themselves.
- It is the process of ensuring individual of a piece of software the code level are functional and work as they were designed to.

3.1.2 System Testing

- System testing is a black box testing method used to evaluate the completed and integrated system, as a whole, to ensure it meets specified requirements.
- The functionality of the software is tested from end-to-end and is typically conducted by a separate testing team than the development team before the product is pushed into production.

3.2 Result and performance Analysis

- As our project is movie recommendation system. One can develop a movie recommendation system using content based or collaborative filtering or a combination of both.
- Advantages: It is easy to design and it takes less time to calculate.
- Disadvantages: The model can only make recommendations based on the user's existing interests.

3.3 Comparison(Table)

References	Methodology	Dataset	Accuracy
K.Sparck Jones	Water Fall	MovieLens Dataset	97.007
Pradeep NK	Neural Network Model	IMDb Datasets	98.080
M. J. McGill	CNN Model	Kaggle Datasets	87.00201

Table 3.1: Table of Comparison

Chapter 4

Conclusion

4.1 Discussion

- In this project, we trying to improve the accuracy, quality and scalability of movie recommendation system, a Hybrid approach by unifying content based filtering Cosine Similarity is presented in the proposed methodology.
- It is beyond doubt that the further development potential of this recommender system is great and by seizing this opportunity, when it will be completed and released, it could play a considerable role in the future.
- Information retrieval has become very difficult nowadays because of the overloading of data and this issue has restricted the users from accessing the items that best match their preferences. This is where the recommendation system comes into use.
- It helps to retrieve personalized information for different users. Also, movies have become a popular medium of entertainment but also the concern of what to watch is also rising along.
- The content-based recommendation approach does not consider other user profiles while making recommendations.
- This will help the user get personalized suggestions for their input.

4.2 Limitations

- While content-based movie recommendation systems have their strengths, they also come with certain limitations.
- Here are some common limitations associated with content-based recommendation systems:
- Limited Diversity in Recommendations.

- Cold Start Problem.
- Dependency on Feature Quality.
- Limited Serendipity.
- Difficulty Handling Changing Tastes.
- Limited Exploration.

4.3 Scope of Future Work

- Enhanced Personalization.
- Integration of Multiple Data Sources.
- Improved Feature Extraction.
- Real-Time Updates.
- Contextual Recommendations.

References

- 1 Pradeep NK, Rao Mangalore KK, Rajpal B, Prasad N, Shastri R. Content based movie recommendation system. International journal of research in industrialengineering. 2020 Dec 1;9(4):337-48.
- 2 K. Sparck Jones. "A statistical interpretation of term specificity and its application in retrieval". Journal of Documentation, 28 (1). 1972.
- 3 G. Salton and M. J. McGill. "Introduction to modern information retrieval". 1983.