

NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY,
BELGAUM, APPROVED BY AICTE & GOVT.OF KARNATAKA)



PROJECT REPORT

on

MOVIE RECOMMENDATION SYSTEM USING MACHINE LEARNING

Submitted in partial fulfilment of the requirement for the award of Degree of

Bachelor of Engineering

in

Computer Science and Engineering

Submitted by:

KARTHIK G	1NT17CS082
KISHORE KUMAR A R	1NT17CS094
K V DHARMA TEJA	1NT17CS095
NAGARAJU P	1NT17CS110

Under the Guidance of

Dr. Shiva Darshan S L
Asst. Professor, Dept. of CS&E, NMIT



Department of Computer Science and Engineering

(Accredited by NBA Tier-1)

2020-2021

NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM

, APPROVED BY AICTE & GOVT.OF KARNATAKA)

Department of Computer Science and Engineering

(Accredited by NBA Tier-1)



CERTIFICATE

This is to certify that the **Movie Recommendation System** is an authentic work carried out by **Karthik G. (INT17CS082), Kishore Kumar A R. (INT17CS094), K V Dharma Teja. (INT17CS095) and Nagaraju P. (INT17CS110)** bonafide students of **Nitte Meenakshi Institute of Technology**, Bangalore in partial fulfilment for the award of the degree of **Bachelor of Engineering** in COMPUTER SCIENCE AND ENGINEERING of Visvesvaraya Technological University, Belgavi during the academic year **2020-21**. It is certified that all corrections and suggestions indicated during the internal assessment has been incorporated in the report. This project has been approved as it satisfies the academic requirement in respect of project work presented for the said degree.

Internal Guide

Dr. Shiva Darshan S L.
Asst. Professor, Dept. CSE,
NMIT Bangalore

Signature of the HOD

Dr. Thippeswamy M. N.
Professor, Head, Dept. CSE,
NMIT Bangalore

Signature of Principal

Dr. H. C. Nagaraj
Principal,
NMIT, Bangalore

Signature of Examiner (Internals)

1.

2.

Signature of Examiner (SEE)


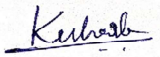

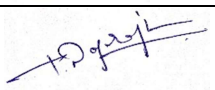
1.

2.

DECLARATION

We hereby declare that

- (i) The project work is our original work
- (ii) This Project work has not been submitted for the award of any degree or examination at any other university/College/Institute.
- (iii) This Project Work does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
- (iv) This Project Work does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a) their words have been re-written but the general information attributed to them has been referenced;
 - b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.
- (v) This Project Work does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

NAM E	USN	Signature
Karthik G	1NT17CS082	
Kishore Kumar A R	1NT17CS094	
K V Dharma Teja	1NT17CS095	
Nagaraju P	1NT17CS110	

Date:


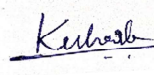

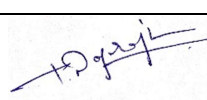
ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crowned our effort with success. We express our sincere gratitude to our Principal **Dr. H. C. Nagaraj**, Nitte Meenakshi Institute of Technology for providing facilities.

We wish to thank our HoD, **Dr. Thippeswamy M.N.** for the excellent environment created to further educational growth in our college. We also thank him for the invaluable guidance provided which has helped in the creation of a better project.

We hereby like to thank our guide **Dr. Shiva Darshan S L, Asst. Professor**, Department of Computer Science & Engineering on her periodic inspection, time to time evaluation of the project and help to bring the project to the present form.

Thanks to our Departmental Project coordinators. We also thank all our friends, teaching and non-teaching staff at NMIT, Bangalore, for all the direct and indirect help provided in the completion of the project.

NAME	USN	Signature
Karthik G	1NT17CS082	
Kishore Kumar A R	1NT17CS094	
K V Dharma Teja	1NT17CS095	
Nagaraju P	1NT17CS110	

Date:

ABSTRACT

Currently, movie recommendation systems have become predominant intelligent systems which play a crucial role in providing selective information. Conventional approaches in recommender solutions are collaborative filtering, and content-based filtering has certain constraints when involving multiple domains. These existing approaches prerequisites user history and emotions to help the user find the desired. In order to address these issues, this paper proposes a movie recommendation system using sentiment analysis and displays the movie comment review in one word along with emoji by analyzing reviews from the user. Comparative analysis was made considering various machine learning algorithms to find the better one considering the certain standard performance metrics. Experimental results showed that the proposed sentiment-based model illustrates better performance using the Naive Bayes model as compared to the traditional approach.

TABLE OF CONTENTS

(Pg.no)

CHAPTER 1: INTRODUCTION	8
1.1 Background	9
1.2 Brief History of Concept	10
1.3 Applications	12
1.4 Research motivation and Problem statement	
1.4.1 Research Motivation	12
1.4.2 Problem Statement	13
1.5 Research objectives and contributions	
1.5.1 Primary Objectives	13
1.5.2 Main Contributions	13
1.6 Organization of the Report	14
1.7 Summary	14
 CHAPTER 2: LITERATURE SURVEY	 15
2.1 Introduction	15
2.2 Related work	15
2.3 Study of Tools or Technology	19
2.4 Summary	21
 CHAPTER 3: SYSTEM REQUIREMENTS SPECIFICATIONS	 22
3.1 General Description	
3.1.1 Product Perspective	22
3.2 System Requirements	
3.2.1 Hardware Requirements	22
3.2.2 Software Requirements	23
3.2.2.1 Functional Requirements & Non-functional Requirements	
3.2.2.2 User Requirements	
3.3 Summary	24

CHAPTER 4: DESIGN	25
4.1 Architectural Design	25
4.2 Data Flow Diagram	27
4.3 Summary	28
 CHAPTER 5: IMPLEMENTATION	 29
5.1 Description of Process	29
5.2 Pseudo-code	30
5.3 Summary	32
 CHAPTER 6: RESULTS	 33
6.1 Experiment results	33
6.2 Test Cases	36
 CHAPTER 7: IMPACT OF PROJECT TOWARDS SOCIETY	 37
 CHAPTER 8: CONCLUSION	 38
 CHAPTER 9: REFERENCES	 39
 CHAPTER 10:SELF ASSESSMENT OF PO-PSO ATTAINMENT	 43
 APPENDIX 1 : PAPER SUBMITTED	
 APPENDIX 2: PLAGIARISM REPORT	

CHAPTER 1

INTRODUCTION

The Internet has become the key part of our lives. In this present world, users frequently come across information overload which is a problem. There are many domains that use recommendation systems. MRS helps a user to get a movie of his interest. Movies are one of the major sources of entertainment that provide relief in our day to day hectic schedules. The recommendation systems are quite often used in video streaming apps and shopping platforms [1]. Although there exists many types of recommendation algorithms, the collaborative filtering algorithm is a popular algorithm. The fundamental purpose is to utilize the inclinations of a group with indistinguishable interests to counsel based on users interest. Users provide a significant extent of response (such as rating) consisting of both positive and negative information to accomplish the need for filtering that helps in giving filtered information. Existing conventional user interest-based recommendation methods are hard to express the requisite information of the data and necessitates manual extraction of features, which consume a lot of time and energy on data labelling and processing [2]. The extraction of the most significant features often determines the efficiency of the algorithm. In recent years, machine learning or deep learning techniques are favoured by researchers as these techniques can accurately express more necessary data information. For instance, the existing recommendation model gets the user ratings to the movie, based on which movies get classified and then recommended to the users. But nowadays users have started giving comment reviews. If a user needs to know about a film, then they have to read the comment completely. Not Everyone has the patience to read the whole of the reviews.

In order to address this issue, this paper proposes a MRS using sentiment analysis that shows movies comment review in one word along with emoji by analyzing reviews from the user. It also recommends movies to the user based upon the movie input provided by the user. Firstly, the user needs to enter the name of the movie about which he wants to know the details. The user can read the short description about the movie and also get to know various details like year of release, main actors involved, etc. The user can also learn about a particular actor in the film upon clicking on their picture. A list of similar movies will also be shown to the user with the help of sentimental analysis [3]. Sentimental analysis is done to understand the emotions, opinions of the people. It is a method of classifying the reviews into either “GOOD” or “BAD”. If the user likes the movie then they can give a positive reaction using smiley emoji

and if they do not like the movie he can give a sad emoji. Based on this, a set of movies are recommended to the user.

The text reviews given by the user are subjected to sentimental analysis and are assigned polarity scores. The sentimental Analysis of the user reviews helps us to categorize the movies. Polarity points are given to the positive words which occur in the review and based on that score, the review is classified as either positive or negative and we also assign an emoji for the same. We also display the details of the movie in different regional languages apart from English. The trailer of the movie is also available in the description. The key objectives of the proposed MRS are as follows:

- To develop a MRS with improved efficiency to recommend the movies, better than the existing model.
- To design a MRS that provides a mechanism that helps users in categorizing movies with similar interests.
- Implementation of MRS using sentiment analysis to classify the reviews and recommend movies.
- Analysis of the performance of the proposed MRS using standard performance metrics.

1.1 Background

MRS is currently one of the leading research areas. Because of the swift progress in Internet technology, the present community has set foot in the period of Web 2, where information overload has turned into materiality. Video streaming services are more often used by customers to access movie content. Movies are the major source of entertainment but that also comes with the problem of information overload. The proposed MRS focuses on introducing several topics related to recommendation systems. The mission of Recommender System is to provide win-win situations for both customers and content providers.

1.2 Brief history of Technology/concept

Machine learning

This is especially useful in healthcare since there is so much data, and if it is carefully given to an intelligent system and trained, the resultant prediction model will be unrivalled, free of human mistakes, and cut diagnosis time in half. ML is a part of AI that allows computers and computer networks to learn and develop without being explicitly taught by humans. Machine learning is based on the creation of computational programmes that can collect data and learn on their own[23][25]. This is especially beneficial in healthcare since there is so much data, and if it is correctly fed and educated into an intelligent system, the resultant prediction model will be unmatched, free of human mistakes, and minimise diagnosis time.

Sentimental Analysis

The primary job in sentimental analysis is to classify the polarity of a given review/text in a paragraph or sentence as whether the expressed feature of the sentence is positive, negative, or neutral by using Natural Language Processing (NLP) and text analysis. Sentimental analysis proved itself to be a beneficial approach for recommender systems [7][10]. A recommender system focuses on anticipating the choice for an item of the target user. Conventional recommender systems work on exact data sets. Nowadays, users provide their feedback, comment or text reviews to the items in many ecommerce or social networking services [17]. These users provide feedback or text data containing numerous sentimental opinions towards the item. This drew out feature describes the same as metadata in content-based filtering. For sentiment analysis, Natural Language Processing Toolkit (NLPTK) is a package in python that represents the output which shows each text review or sentence is positive, negative or neutral [9]. To compute at a high pace, we have to keep the set of words, and then classify the expression as “GOOD” or “BAD”, which requires a lot of time. The challenge comes when the data is from a language other than English. We cannot rely on changes which will be inconsistent [10][23].

Cosine Similarity

To provide accurate recommendation and matching to the users, need a method to define similarity measures. There are some methods: Cosine similarity, Euclidean distance, Pearson coefficient and others. Cosine similarity, which is a very commonly used approach for recommendations, has an excellent effect of similarity and a reasonably high accuracy among the rest. This technique is frequently used in NLP and text information processing. It measures similarity between the two n-dimensional vectors which are roughly pointing in the same direction by using the angle between them[11][12] .

$$\cos \theta = \frac{\underline{a} \cdot \underline{b}}{||\underline{a}|| ||\underline{b}||} \quad (1)$$

$$||\underline{a}|| = \sqrt{a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2} \quad (2)$$

$$||\underline{b}|| = \sqrt{b_1^2 + b_2^2 + b_3^2 + \dots + b_n^2} \quad (3)$$

The ‘.’ symbol in Eq 1 represents the dot product between vectors, and Eq 2, Eq 3 represents the length between two vectors, which is, the square root itself. Because the cosine of 0 is 1 and any angle between (0,) is less than 1, the value of the cosine similarity is an integer between 0 and 1. If the value is close to 1, two are comparable; otherwise, two are not similar[11].

Cosine similarity is most commonly used in the information retrieval and text mining. For example, to find how similar the two documents are, based on how many times the terms or words in the document appear, cosine similarity calculates whether the documents are similar or not. It does not takes difference between average scores of items, it takes original cosine similarity values which solves the problem effectively.

1.3 Applications of Recommender Systems

Recommender systems are information retrieval technologies that provide individual users with accurate recommendations. People can access their favourite movies with the use of collaborative movie recommender systems, which suggest comparable users or movies based on their previous common ratings.

A recommender system attempts to anticipate or filter preferences based on the user's preferences. Recommender systems are used in a wide range of applications, including movies, music, news, books, research articles, search queries, social tagging, and general items.

Recommender systems have the advantage of providing personalisation for e-commerce customers, hence boosting one-to-one marketing. As part of their marketing strategy, Amazon, a pioneer in the usage of collaborative recommender systems, offers "a personalised store for every customer."

Personalization for e-commerce customers is a benefit of recommender systems, which boosts one-to-one marketing. Amazon, a pioneer in the use of collaborative recommender systems, offers "a customised store for every customer" as part of their marketing strategy.

Recommendation systems have the ability to transform the way websites communicate with their users, as well as assist businesses to maximise their ROI by gathering data on each customer's preferences and purchases.

1.4 Research motivation and Problem statement

1.4.1 Research Motivation

- In this fast-paced world, we all need entertainment to keep our spirits up and our energy levels up.
- In today's hectic society, recommendation systems are becoming increasingly vital.
- People are always pressed for time due to the numerous tasks they must complete in the allotted 24 hours.
- The purpose of a recommendation system is to find stuff that would be engaging and intriguing to an individual.
- It also has a number of features that allow each user/individual to construct customised lists of useful and fascinating stuff.

- As a result, the movie recommendation system is critical because it assists users in making the best decisions without requiring them to squander their cognition.

1.4.2 Problem Statement

" Design and development of a movie recommendation system that analyzes the users' behaviour and suggests movies to the users based on similar tastes and interests."

1.5 Research objectives and contributions

1.5.1 Primary objectives

- To develop a movie recommendation system with the improved efficiency to recommend the movie better than the existing model.
- To design a movie recommendation system that provides a mechanism to assist users to group people who have similar interests together.
- Implementation of movie recommendation system incorporating sentimental analysis to classify the reviews.
- Analysis of the performance of the proposed movie recommendation system using standard performance metrics.

1.5.2 Main contributions

- We will be doing the sentimental analysis of the reviews, which gives only two possibilities i.e., either "Good" or "Bad".
- We built an user interface which helps the user to interact easily with the system
- We will also deploy this product in the cloud for better performance, easy access and availability.

1.6 Organization of the report

The remainder of this paper is organized as follows: [Chapter 1](#) gives the brief introduction and the background study of the model of movie recommendation system which includes problem statement, objectives and contributions of project. Chapter 2 describes literature survey and related work of project. [Chapter 3](#) gives detail about the requirements such as functional, non functional and user requirements. Then we expounded the design and description part of the project along with its pseudocode in [Chapter 4,5](#). In [Chapter 6,7](#) Test cases and experiment results are described. Finally, we explained the impact of our project towards society and concluded it in further chapters.

1.7 Summary

- ✓ Creating a website to recommend movies to the users.
- ✓ Using sentimental analysis for movie recommendation.
- ✓ Displaying the movie comment review in one word along with emoji.
- ✓ Adding reviews from the user.

CHAPTER 2

LITERATURE SURVEY

2.1 Introduction

The Internet has become a key part of our lives. In this present world, Information overload is a common concern for users. There are many domains that use recommendation systems. Movie Recommendation System helps a user to get a movie of his interest. Movies are one of the major sources of entertainment that provide relief in our day to day hectic schedules. The recommendation systems are generally used in digital entertainment, such as Netflix, Prime and e-commerce platforms. We designed a website that gives data about many movies. It also recommends a set of movies to the user based on the movie input provided by the user.

2.2 Related work

Many proposed Recommendation Systems have existed over the decades. These systems employ various methods, like Collaborative, Content-Based, Hybrid Filtering, sentiment analysis, to suggest the preferred items. These existing methods have special attention in Recommendation Systems, and each of them are discussed, see [Table 1](#)

Table 1 : Details related to the literature survey.

Sl.no	Method	Description	References
[1]	Collaborative Filtering	The goal of a recommendation system is to forecast which items a user will be interested in based on their preferences.	Mukesh Kumar Kharita; Atul Kumar Pardeep Singh
[2]	Multinomial Naïve Bayes	The input dataset is made up of a collection of tweets that convey the opinions of various people about the various election candidates who are competing in the election.	Prabha PM Surya Lakshmi V Seetha; B Subbulakshmi

[3]	Semi-Supervised learning approaches	The recommendation system or suggestion engine discussed in this paper assists a person in identifying noteworthy and potentially relevant services or products depending on his or her interests.	Sushmita Roy; Mahendra Sharma; Santosh Kumar Singh.
[4]	Support Vector Machine	Based on the TMDB dataset, a sentiment analysis system is constructed that uses SVM to estimate favourable and unfavourable sentiment from user reviews, and then evaluates movies based on sentiment analysis and review rating.	Nimish Kapoor; Saurav Vishal; Krishnaveni. K.S
[5]	Multinomial and Bernoulli Naïve Bayes	To forecast whether the sentiment of the news storey is good or negative, the two most common techniques to Naive Bayes Text Categorization, Multivariate Bernoulli Naive Bayes Classification and Multinomial Naive Bayes Classification, are employed.	Gurinder Singh; Bhawna Kumar; Loveleen Gaur; Akriti Tyagi
[6]	K-Means and TF-IDF	A cluster or group, which is a collection of tweets made by various users on the Twitter website, was constructed and the user was informed about the behaviour, which is mostly focused on sentimental analysis.	K Sai Madhu; B Chakradhar Reddy; CH Damarukanadhan

[7]	Cosine Similarity	A method for extracting the core material from web documents is proposed. The Content Structure Tree is the basis for the algorithm (CST). To begin, the suggested system employs HTML Parser to generate DOM trees from which CST can be generated.	Swe Swe Nyein.
[8]	Hybrid filtering	The goal of using movie tweets is to gain a better understanding of present trends, people's opinion, and customer reaction to the film. Experiments using the public database have generated encouraging results.	Sudhnshu Kumar; Kanjr De; Prtha Pratim Roy.
[9]	Naïve Bayes	On the basis of student input, a few learning algorithms were used to successfully grade the teachers of an educational institute.	K. S. Kishnaveni; Rohit R Pai; Vignesh Iyer.
[10]	Deep learning method	Proposed a deep learning strategy based on autoencoders to create a collaborative filtering system that predicts a user's movie ratings based on a big database of other users' ratings.	Jeffrey Lund; Yiu-Kai Ng.
[11]	Cosine similarity	The directed graph is built using the grants-in-aid for scientific research database, with three types of edges created based on research partnership history, cosine similarity of research contents, and	Kanta Nakamura; Kazushi Okamoto

		their combination.	
[12]	Recommendation System	Examines the results of recommendations obtained with and without taking into account movies that have never received an above-average rating, where average rating is defined as the midpoint between 0 and the maximum rating, for example, 2.5 on a 1 to 5 rating scale.	Muppana Mahesh Reddy; R. Sujithra Kanmani; B. Surendiran.
[13]	Similarity Analysis	The author proposes a method for comparing movie screenplays to score similarities between them. To analyse, the approach we provide does not require any rating information or metadata.	Sung Min Kim; Young Guk Ha
[14]	Collaborative approach	Developed a collaborative recommender system that matches students' course histories to recommend online courses to them. To find patterns between courses, the suggested method uses data mining techniques.	Raghad Obeidat; Rehab Duwairi; Ahmad Al-Aiad

2.3 Study of Tools/Technology

Programming Language - Implementation

Python generates code that is both succinct and understandable. Machine learning and AI are based on complicated algorithms and flexible workflows, while Python's simplicity allows developers to create trustworthy solutions. Instead of focusing on the technical shading

of the language, developers may devote all of their attention to addressing an ML problem. Python is also appealing to many developers since it is simple to learn. Humans can understand Python code, making it easier to create machine learning models. Python, according to many programmers, is more intuitive than other programming languages. Others point to the numerous frameworks, libraries, and extensions that make implementing certain functionality easier. Python is often regarded as a good choice for collaborative development when several developers are involved. Python is a general-purpose language that can do a variety of difficult machine learning tasks while also allowing you to rapidly create prototypes.

Extensive selection of libraries

A large number of libraries and frameworks are available. Implementing AI and machine learning algorithms may be difficult and time-consuming. To allow developers to come up with the finest coding solutions, they need a well-structured and well-tested environment. Programmers use a variety of Python frameworks and modules to reduce development time. A software library is a collection of pre-written code that programmers may utilise to tackle common programming challenges. Python contains a large-scale artificial intelligence and machine learning library. Here are a few examples:

- Machine learning frameworks include Keras, TensorFlow, and Scikit-learn.
- NumPy is a Python package for scientific computing and data analysis.
- SciPy is a Python package for advanced computation.
- Pandas is a data analysis tool that may be used for a variety of purposes.
- Seaborn is a data visualisation platform.
- Scikit-learn has a number of classifications and is meant to interact with the NumPy and SciPy Python numerical and scientific libraries.

Text Mining

The NLTK (Natural Language Toolkit) package, which includes the Splitter class [7][11], is used in text mining. Tokenization is the process of breaking a paragraph into multiple sets of sentences or separating sentences into various sets of words and then storing the results in another list. For example, consider a review “The movie was Great !” The statement is tokenized and kept in the list as shown.

```
['The', 'movie', 'was', 'Great', '!']
```

Once the tokenization is done the second step would be cleaning the data which means removing all the special characters. and the third step is removing stop words that aren't useful in the analytics section. like he/she, the, is, etc.

Next we need to check for emotions [26], so the list is fed to the Naive Bayes model to classify the words whether it is a positive word or negative word. For a positive word we give a sentiment score as +1, and for a negative word we give -1. Emotional weight is adjusted throughout time, and the more data collected, the more exact the weight becomes. [7] [8][10]. After all these steps, in the above said case, the list is made to keep track of emotions is ['Great']. So the overall sentimental score of sentence is 1.

Multinomial Naive Bayes

The text data which is already mined is fed to the Naive Bayes model that categorizes the text word into either positive and negative. The term frequency of a term in a text or sentence is determined using the multinomial model [13]. This is useful in sentimental analysis of a particular term, resulting in an effective decision. Term occurrence or frequency aids in determining if a term is helpful in analysis or not. However, if a word or term appears multiple times in a document, it could be a stopword that has no bearing on the review or sentence, such words or terms must be eliminated to get better results by using Naive Bayes algorithm [13][14].

Cosine similarity

To provide accurate recommendation and matching to the users, we need a method to define similarity measures [18]. There are some methods like Cosine similarity, Euclidean distance, Pearson coefficient and others. Cosine similarity is one of them, and it provides a solid similarity measure with a relatively high precision, which is a very regularly used technique for recommendations. This technique is frequently used in text information processing and natural language processing [19]. In [Fig 1](#) we can see, It measures similarity between the two n-dimensional vectors which are roughly pointing in the same direction by using the angle between them.

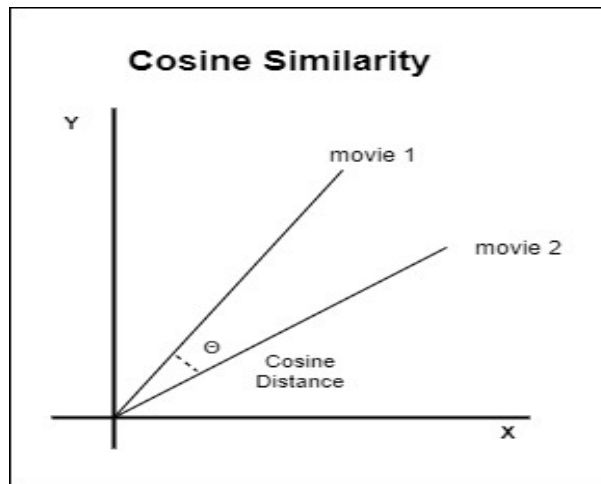


Fig. 1: Cosine Similarity

2.4 Summary

We designed a website that gives data about many movies. It also suggests a collection of films to the user based on the movies that the user has selected. Firstly, the user needs to enter the name of the movie about which he wants to know the details. The user can read the short description about the movie and also get to know various details like year of release, main actors involved. The user can also learn about a particular actor in the film upon clicking on their picture. A list of similar movies will also be shown to the user with the help of sentimental analysis.

CHAPTER 3

SYSTEM REQUIREMENTS SPECIFICATIONS

This chapter deals with the system requirements part. The different types of requirements like hardware requirements, Software requirements, functional requirements and user requirements are being discussed which are used in our project. General description is also discussed which emphasises on product perspective.

3.1 General Description

3.1.1 Product Perspective

The proposed solution is designed to:

- We use content based filtering for an individual pertaining to there interest
- With the help of the above model we build a recommendation system which gives higher accuracy, performance and better user experience
- We also use cosine similarity for sentiment analysis of the reviews from the users

3.2 System Requirements

3.2.1 Hardware Requirements

Table 2 : Hardware requirements with description.

Sl. No.	Hardware Requirements	Description
1	Processor	1.6 GHz CPU
2	Hard Disk	4 x 20 GB of free space or more
3	RAM	4/8 GB

The sort of requirements defined by software application and OS are physical resources of computer or hardware. Some of the hardware required for our project is interpreted, see [Table 2](#).

3.2.2 Software Requirements

Table 3 : Software Requirements with description

Sl. No.	Software Requirements	Description
1	Operating system	Windows-10/ Windows-7/Windows-XP
2	Software Tool	Anaconda/Jupyter
3	Coding Language	Python, HTML,JS,CSS

Software requirements specify the software resources of computers that are required to be installed on the device to yield the best performance of an application. Some of the software requirements required for our project are discussed, see [Table 3](#).

3.2.2.1 Functional Requirements & Non-functional Requirements

Functional Requirements

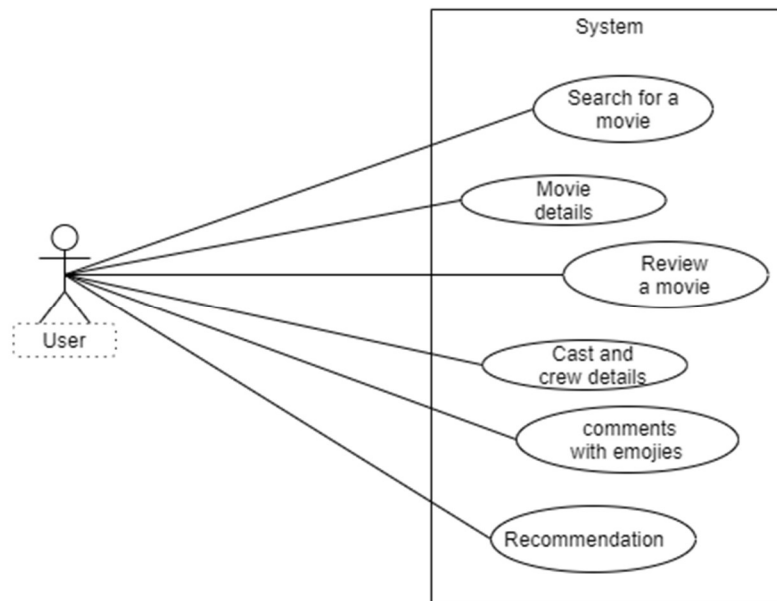


Fig. 2: Use Case Diagram

[Fig. 2](#). shows the Use-case diagram representing the functionalities of the user. The user can search for a movie in the search engine, he can view the movie details, he can review a movie,

he can view the cast details, he can see the sentiment analysis result of the reviews and can also see the list of recommended movies.

Non-functional Requirements

- **Usability** : The model should have a clean interface with simple design that is easily usable.
- **Availability** : The recommendation made will be accurate 24X7 and the content will be accessible from anywhere at any time.
- **Scalability** : However large the dataset might be, it gives an accurate prediction.
- **Performance** : This project has very high performance even though datasets are very huge.
- **Supportability** : The project will run on Windows, Linux as well as other major operating systems.
- **Cost effective** : As we use no hardware devices we are producing a cost effective model.

3.2.2.2 User Requirements

1. **User Interface**: since the user interacts with the application through an interface, an attractive and simple to use User Friendly Interface is required to be designed.
2. **Means of Access**: All the user needs to access the application is a smart device and a stable internet connection.
3. **Account**: The user has to register themselves and then login with userID and password. Only after logging in can the user access his account information. This provides Security to user data.
4. **Identity**: user has to verify his identity (mail/mobile OTP) before gaining complete access to application.

3.3 Summary

In this chapter we discussed the requirements part that was used in our project. Starting with the product perspective which tells how the proposed solution should be. Hardware requirements which were majorly system components along with their configurations. Software requirements like OS, language which we used is being mentioned. Functional requirements are displayed with Use-case diagram. Non-functional requirements include the general performance measures and User requirements are also discussed.

CHAPTER 4

DESIGN

This chapter deals with the architectural design of our proposed work. The components that are required for our work are shown in Fig3. This architecture diagram gives us the overview of our project.

4.1 Architectural Design of the Proposed Work

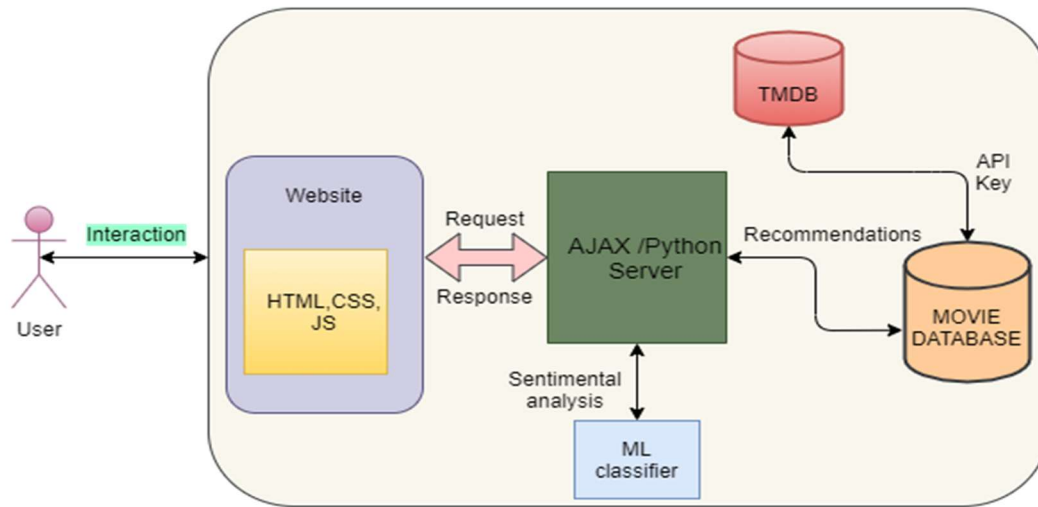


Fig. 3: Architecture Diagram

The general overview of the architectural framework of the proposed MRS using sentiment to classify reviews and recommend movies with similar interests, see Fig.3.

Database Description

There are public databases available which are employed to suggest movies and other forms of entertainment media¹. In our proposed work, the Movie Database (TMDB) Application Programming Interface (API) was used to fetch more features of the movies by requesting the movie data from TMDB using API key. Upon processing it sends particular movies information along with all the information received from TMDB. After receiving all the necessary information from the TMDB, the Movie Database contains all the below mentioned attributes to be displayed further in the website, see Table 4.

Table 4: Details related to the movie database

Sl. No.	Attributes	Sl. No.	Attributes
1	Movie Title	6	Movie Release Date
2	Movie Poster	7	Movie Runtime
3	Movie Overview	8	Movie Status
4	Movie Ratings	9	Movie Trailer
5	Movie Genre	10	Movie Cast and Crew Information

The TMDB is a database that collects various forms of information about films, television shows, actors, and so on. It has an API that allows the developers to gather information regarding various movie genres. We mostly focused on film data. The TMDB API provides a number of endpoints for collecting data.

We looked at the dataset to gain insight into the movie dataset, which will help us create the proposed MRS utilising Python modules. The number of movies in each genre, the most rated genres, the number of movies rated in each rating group, and the most rated movies were all discovered.

Interface Development

We build the website using HyperText Markup Language (HTML) for the skeleton, Cascading Style Sheets (CSS) for stylings of the webpage, and JavaScript (JS) for different functionalities to provide an easy interface of the application to the user. We have hosted the website using port no.:5000 (localhost:5000) and accessed it on the local system to demonstrate the proof of concept of the proposed MRS. In the User Interface (UI), we employ a search engine to search any particular movie of the user choice, and upon clicking the search button a request to Asynchronous JavaScript And XML (AJAX) server is sent as an argument. On receiving the response from AJAX server various details like title, movie poster, overview, ratings, genre, release date, runtime, status, movie trailer, and cast and crew information are displayed. The proposed MRS employs sentiment analysis and displays the movie comment review in one word along with emoji by analyzing reviews from the user. The provision is also made to the user in the website to view the details of the recommended movies in any regional language the user requires.

AJAX Server

AJAX is a set of web development approaches using many web technologies on the client's end to create asynchronous web applications. We will be fetching movie data from a free API and displaying them to the user. There are many API available for movie data on the internet. The title of the movie is sent to the Machine Learning (ML) Classifier to perform sentimental analysis and then a request is sent to the movie database to get the particular movie information and also the recommendation.

ML Classifier

The sentiment in reviews is a paramount aspect in evaluating a movie. It is basically used for data processing to recommend accurate movies to the users. In the proposed MRS, the ML model utilizes the sentimental analysis of text reviews for Classifying the reviews as good or negative and establishing the user's preference for a certain film.

Recommendations

Collaborative filtering is a common method followed in the recommendation system. If the movie recommendations are based on the ratings or genres stated by other users it is quite difficult to provide precise results. To make it easy for users and to overcome this drawback, we used another technique for recommendation which is Cosine Similarity. This method helps us in suggesting precise results to the user as the ratings and genres are given by the user himself and depends completely on the user's choice. Previous methods lacked the accuracy of true recommendations as they were dependent on other users.

4.2 Data Flow Diagram

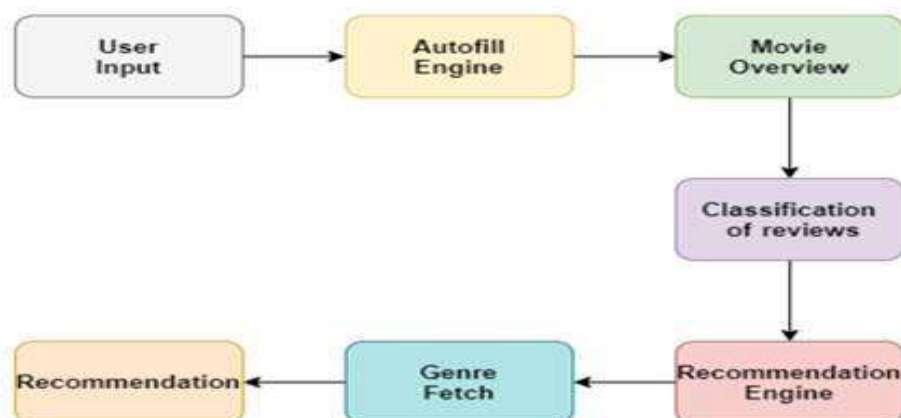


Fig. 4: Data Flow Diagram

[Fig4.](#) shows the flow of the data. The user needs to give an input in the search engine in the user interface. The search engine has auto-complete functionality. On clicking the search button we can see the details of the movie like overview of the movie along with a poster, runtime, status, release date, genre, title, movie cast information along with their pictures, trailer video of the movie, sentiment analysis of the reviews, list of recommended movies. The genre of the movie is being taken for recommendation of movies using sentiment analysis.

Summary

The architecture of the proposed work contains various components that are shown in [Fig.3](#). Each component is explained in detail. Some components include ML classifier, Ajax server, Interface development etc. The flow diagram of the data is also being depicted in [Fig. 4](#). It shows how the data is transferred from one point to another.

CHAPTER 5

IMPLEMENTATION

In this chapter we discuss the implementation part of our project. The description of the process along with the existing approach and our proposed approach is also being discussed. A small pseudocode is also put up to get a better understanding of our work.

5.1 Description of Process

The Internet has become the key part of our lives. In this present world, people are frequently confronted with the problem of information overload. There are many domains that use recommendation systems. Movie Recommendation System helps a user to get a movie of his interest. Movies are one of the major sources of entertainment that provide relief in our day to day hectic schedules. Digital entertainment sites such as video streaming sites and e-commerce platforms all use recommendation systems.

We designed a website that gives the data about many movies. It also recommends a set of movies to the user based on the movie input provided by the user. Firstly, the user needs to enter the name of the movie about which he wants to know the details. The user can read the short description about the movie and also get to know various details like year of release, main actors involved. The user can also learn about a particular actor in the film upon clicking on their picture. A list of similar movies will also be shown to the user with the help of sentimental analysis.

Sentimental analysis is done to understand the emotions, opinions of the people. It's a system for classifying reviews as favourable or bad. If the user likes the movie then they can give a positive reaction using smiley emoji and if they do not like the movie he can give a sad emoji. Based on this, a set of movies are being recommended to the user.

Existing Approach

Currently the user gives ratings to the movie, based on which the rating of the movies are categorised and then suggested to consumers. But nowadays users have started giving comment reviews. If a user needs to know about a film, then they have to read the comment completely. Not everyone has the patience to read the whole of the reviews.

Proposed Approach

The text reviews given by the user are subjected to sentimental analysis and are assigned polarity scores. The sentiment analysis of the reviews helps us to classify the movies. Polarity scores are assigned to the positive words that occur in the review and based on that score, the review is classified as either positive or negative and we also assign an emoji for the same. We also display the details of the movie in different regional languages apart from English. The trailer of the movie is also available in the description.

The main approaches in this work are :

- Creating a website to recommend movies to the users.
- Using sentimental analysis for movie recommendation.
- Displaying the movie comment review in one word along with emoji.
- Adding reviews from the user.

5.2 Pseudo-code

Recommend.js

```
$(function()
{
    // Button will be disabled until we type anything inside the
    input field
})

function recommend card(e)
{
    // will be invoked when clicking on the recommended movies
}

function load_details(my_api_key,title)
{
    // get the basic details of the movie from the API (based
    on the name of the movie)
}
```

```

function movie_recs(movie_title,movie_id,my_api_key)
{
    // passing the movie name to get the similar movies from
    python's flask
    if(recs=="Sorry! The movie you requested is not in our
    database. Please check the spelling or try with some other
    movies")
        // display none
    else
        // call function
        get_movie_details(movie_id,my_api_key,arr,movie_title);

function get_movie_details(movie_id,my_api_key,arr,movie_title)
{
    // get all the details of the movie using the movie id
    //call
    function
    show_details(movie_details,arr,movie_title,my_api_key,movie_id);
}

function
show_details(movie_details,arr,movie_title,my_api_key,movie_id)
{
    // passing all the details to python's flask for displaying
    and scraping the movie reviews using imdb id
}

function get_individual_cast(movie_cast,my_api_key)
{
    // get the details of individual cast

```

```

    return
    {cast_bdays:cast_bdays,cast_bios:cast_bios,cast_places:cast_places};
}

function get_movie_cast(movie_id,my_api_key)
{
    // getting the details of the cast for the requested movie
    returns{cast_ids:cast_ids,cast_names:cast_names,cast_chars
    :cast_chars,cast_profls:cast_profiles};
}

function get_movie_posters(arr,my_api_key)
{
    // getting posters for all the recommended movies
}

```

The website contains a Search bar which will be disabled at the beginning and only enabled when something has entered in it. On hitting the enter button will get the basic details of the movie based on the movie name using API key. The parameters like *movie_details,arr,movie_title,my_api_key,movie_id* are passed to the python's flask for displaying and scraping the movie reviews using imdb, get the details of the individual cast like *cast_bdays, cast_bios, ,cast_places* and get the posters of all the recommended movies.

5.3 Summary

We discussed the implementation part along with the pseudocode for better understanding. We also mentioned the existing approach and also our proposed approach to get a better comparison. The whole description about the process is also described.

CHAPTER 6:

RESULTS

In this chapter we discuss the outcomes of our project. Few images of our website showing important features are being shown. The possible test cases also have been included. Accuracy graph and ROC curve are also plotted to get better comparisons and views.

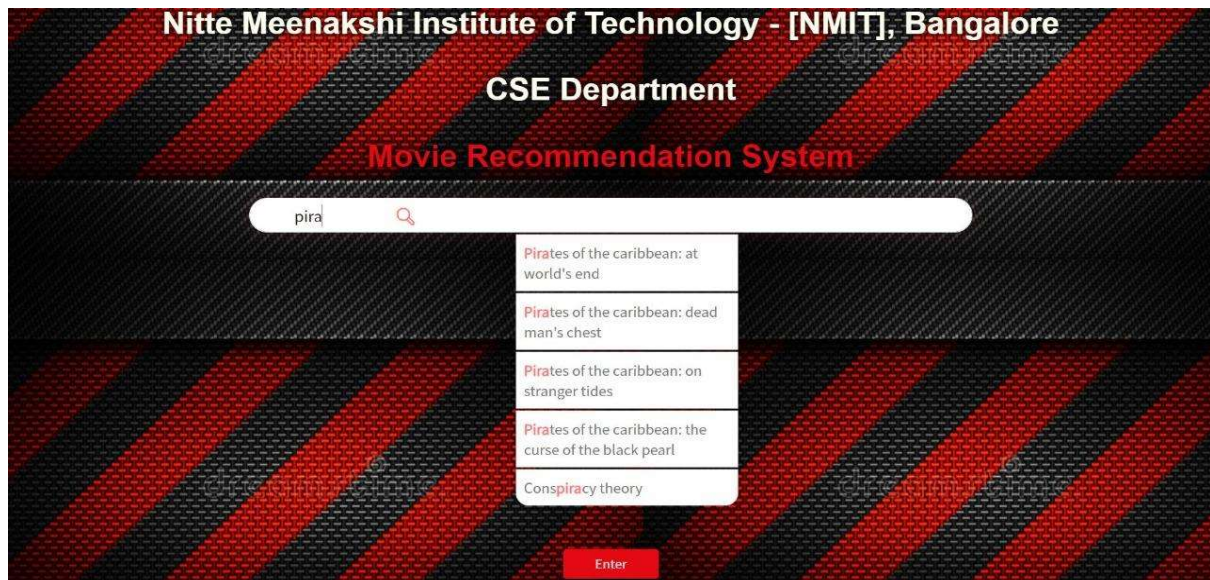


Fig. 5: Auto-fill functionality

USER REVIEWS	
Comments	Sentiments
This third installment is just as long, fixes some of the damages but adds along new ones. At Worlds End improves on its CGI magic but still possesses that dragging story that is just too long and confusing.	Good : 😊
I'm a huge fan of the Pirates of the Caribbean movies, but I was a little nervous about seeing this third installment because of the not as good other third installments. I still had pretty high expectations, and I was not disappointed! The plot was brilliant and very creative, with plenty of exciting twists and turns. As always, the acting was incredible. Johnny Depp always does an amazing job, but his Captain Jack Sparrow is probably one of his best works yet! I still don't see how anyone could be so brilliant... Anyways, overall, just a really great film. Better than the other Pirate films in the trilogy (and you know how good those are!) See it, you won't be disappointed!	Good : 😊
This movie is truly spectacular!!! Non stop action, with a surprising ending. If it weren't for the silliness of the Jack Sparrow clones, I would give it a 9 and 1/2. It takes a little away from the film's momentum. As far as the final battle rivaling that of the LOTR trilogy, I would say no. This trilogy will go down as one of the great trilogies. Breathtaking special effects, and fabulous scenery. All in all, a great film. It must be viewed in the theater. It is better than Dead Man's Chest, but I still like the overall story line of Curse of the Black Pearl better. As always, be sure to stick around after the credits for an extra scene. This one adds to the plot of the film and gives viewers some closure after a heart wrenching finale. Make sure you check this film out.	Good : 😊
Do not listen to the stingy critics!! This movie is AMAZING! Everything you hoped for and more. I can't wait to see it again. The fact that it is a long movie makes it all the better for the viewer because you get more of all the adventure, romance and breathtaking scenes. The writing is incredible, the acting is the best it has been since the beginning, and the artistry is second-to-none. There is no fluff, every scene means something. It's funny, intriguing and keeps you on the edge of your seat. My heart didn't stop pounding until long after the credits ended. It's the kind of movie that stays with you. It's satisfying If this is the last Pirates, and it leaves you wanting more If we are to be blessed with a fourth. Bottom line is: go see this movie, and have a great time doing it!	Bad : 😞
OK. So the first one was a great movie, second one was alright. So I went to the midnight showing of Pirates and accidentally fell asleep during the hour and a half of pointless content that is the first part of the movie. I went and saw it again and watched every minute and let me say, it was so bad and the plot line so choppy that once the movie was over i was unsure of the definitions of reality. Quite literally, i could have made a better	Bad : 😞

Fig. 6: Sentiment Analysis

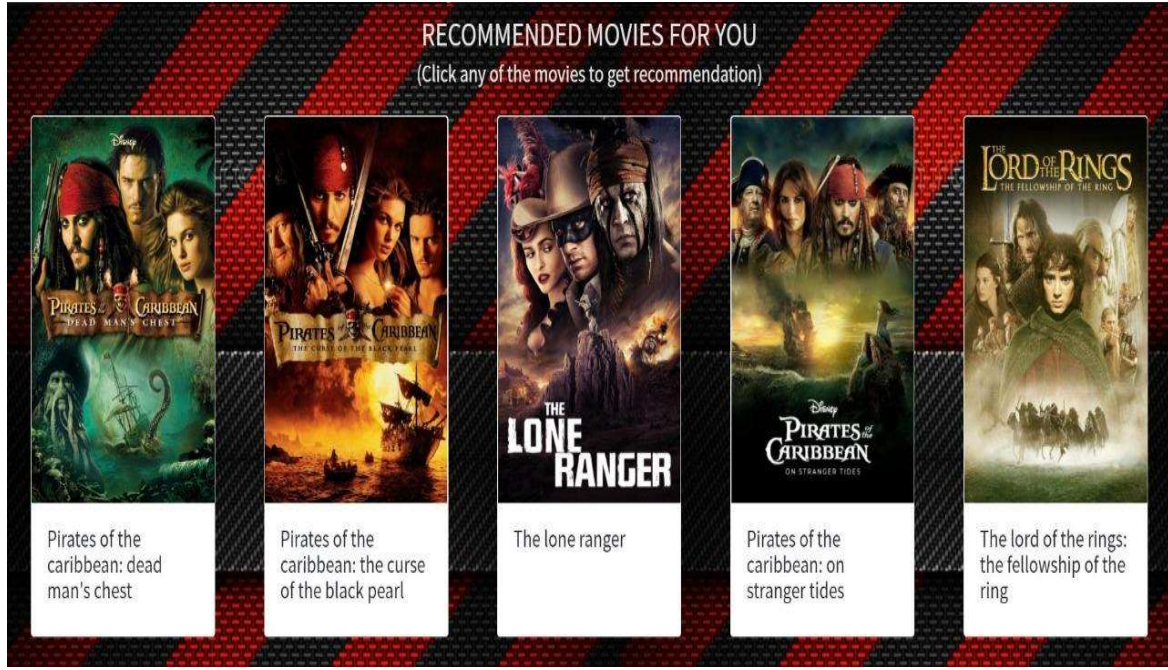


Fig. 7: Recommended Movies

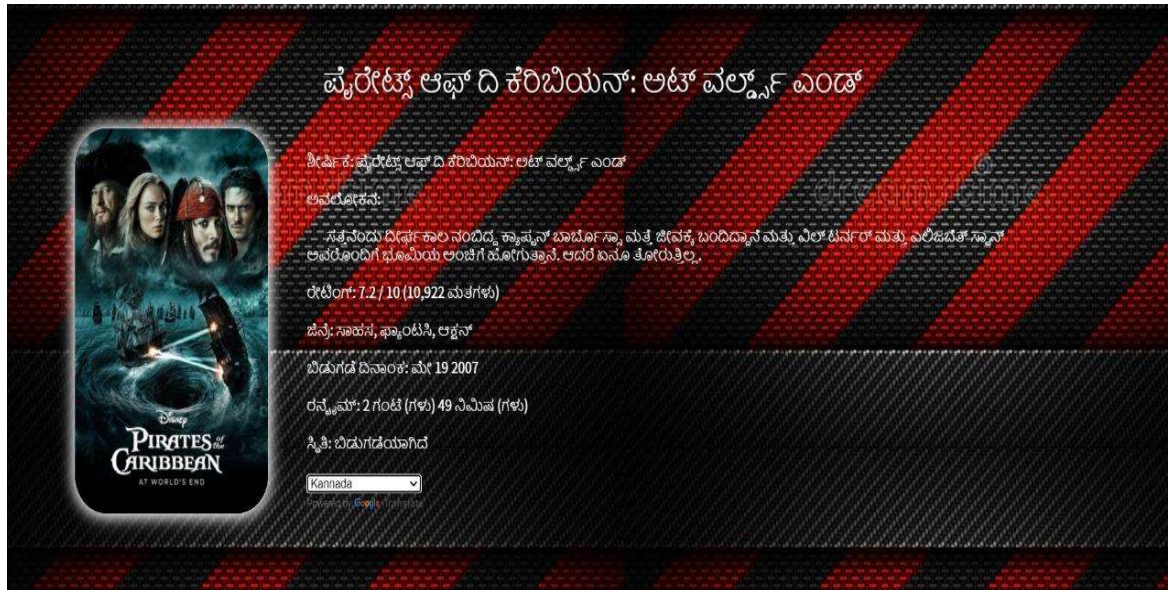


Fig. 8: Regional Language

Fig. 5 shows the auto-complete feature, where we can get a list of movies starting with the letters which we enter. It helps users save time while they are searching for movies with long names. Fig. 6 demonstrates the sentiment analysis part where we analyse the whole review and give one word about the whole review that determines whether the review is good or bad. Fig.

7 shows us a list of movies that are recommended to the user based on the genre using sentiment analysis. Fig. 8 shows the content of the whole page in a regional language i.e Kannada.

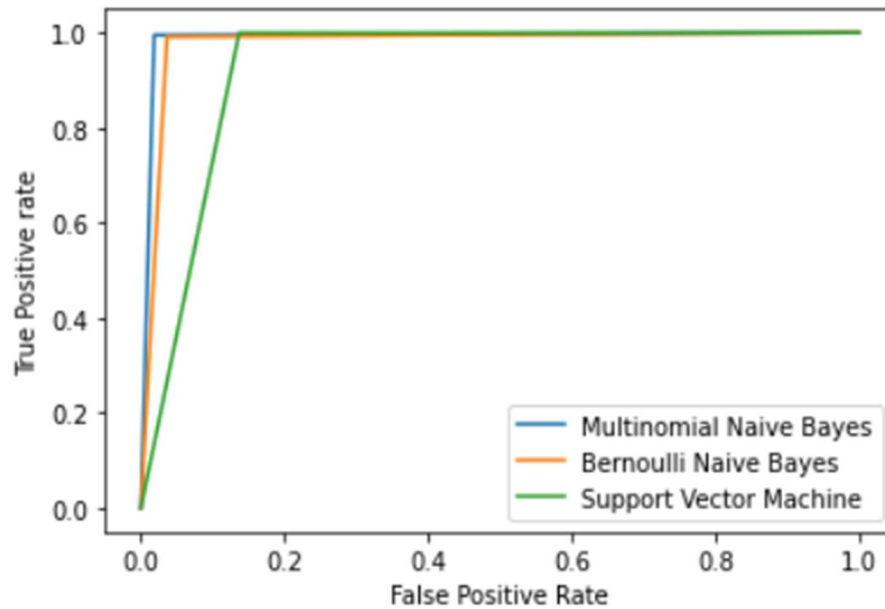


Fig. 9: ROC curve indicating that TPR for Multinomial Naive Bayes algorithm is greater than Bernoulli Naive Bayes and Support Vector Machine ML classifiers.

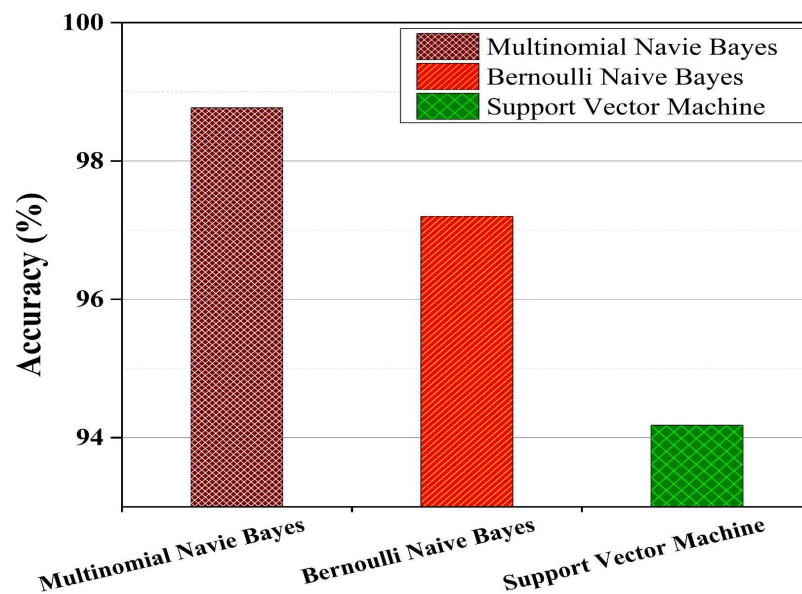


Fig. 10: Comparison analysis of accuracy obtained by various ML classifiers.

As we have observed in Fig. 9 and Fig. 10, The results of Multinomial Naive Bayes on the provided dataset is clearly superior to that of Bernoulli Naive Bayes and the Support Vector Machine classifier, as evidenced by the accuracy bar graph and the ROC curve. Bernoulli Naive

Bayes and Support Vector Machine classifiers achieved an accuracy of 97.20%, and 94.18%, respectively which is less than that of Multinomial Naive Bayes Classifier that is 98.77%. The ROC curve clearly shows that the true positive rate for Multinomial Naive Bayes is higher, indicating that the Multinomial Naive Bayes classifier is superior in terms of performance.

Test Cases :

Table 5 : Test - cases

Sl. No	INPUT	OUTPUT
1.	ahgftyhjf	Sorry! The movie you requested is not in our database. Please check the spelling or try with other movies!
2.	Han	Auto-complete function
3.	Hancock	Overview, rating, genre, release date, runtime, status, top cast, sentimental analysis of reviews, recommended movies.

CHAPTER 7

IMPACT OF YOUR PROJECT TOWARDS SOCIETY/ENVIRONMENT

Movie recommendation algorithms have shown to be the most effective option for dealing with conflicting information. They assist in decision-making by conserving time and energy. Future research will focus on upgrading the current models and techniques used for recommendation systems to increase prediction and recommendation effectiveness. MRS shapes user's choices and helps users with their preferences. These types of recommender systems can be used in other departments like e-commerce websites, healthcare departments etc. Recommender systems help users make right choices in regard to what they want within the stipulated time. People are always in a hurry nowadays. They don't tend to waste time on selecting movies by researching about the movie. That's where MRS plays its part. It helps users with the required options /features from which they can save their precious amount of time. Finally, Movie Recommendation systems can be the best platform in today's world.

CHAPTER 8

CONCLUSION

We've suggested an MRS that recommends movies based on sentiment analysis data. Sentiment analysis provides information on how the viewer reacts to a certain film and how this information may be beneficial. To enhance the suggestions, the suggested method employed a cosine similarity score. Based on our work, the proposed sentiment based MRS exhibits better performance using the Multinomial Naive Bayes classification by accomplishing the accuracy of 98.77% compared to other classifiers. We deduce that our proposed MRS recommends more accurately than the other recommendation systems.

CHAPTER 9

REFERENCES

- [1] Shruti Doshi , Prarthana Bakre , Ruchi Agrawal and Prof. Pranali Kosamkar “Movie Recommender System” ,2016 International Journal of Emerging Science and Engineering (IJESE), 2016.
- [2] Shreya Agrawal and Pooja Jain “An Improved Approach for Movie Recommendation System”,2017 International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017).
- [3] Peng, Xiao, Shao Liangshan, and Li Xiuran. "Improved Collaborative Filtering Algorithm in the Research and Application of Personalized Movie Recommendations", 2013 Fourth International Conference on Intelligent Systems Design and Engineering Applications, 2013.
- [4] Muppana Mahesh Reddy ,Sujithra and Dr. B. Surendiran “Analysis of Movie Recommendation Systems with and without considering the low rated movies” 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE).
- [5] Jeffrey and Yiu-Kai Ng “Movie Recommendations Using the Deep Learning Approach” 2018 IEEE International Conference on Information Reuse and Integration for Data Science.
- [6] Munoz-Organero, Mario, Gustavo A. Ramírez-González, Pedro J. Munoz-Merino, and Carlos Delgado Kloos. "A Collaborative Recommender System Based on SpaceTime Similarities", IEEE Pervasive Computing, 2010.
- [7] K. S. Krishnaveni; Rohit R Pai; Vignesh Iyer,” Faculty rating system based on student feedback using sentimental analysis”, 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI).

- [8] Sudhanshu Kumar; Kanjar De; Partha Pratim Roy, "Movie Recommendation System Using Sentiment Analysis From Microblogging Data Authors", IEEE Transactions on Computational Social Systems.
- [9] Nimish Kapoor; Saurav Vishal; Krishnaveni. K. S., "Movie Recommendation System Using NLP Tools", 2020 5th International Conference on Communication and Electronics Systems (ICCES).
- [10] K Sai Madhu; B Chakradhar Reddy; CH Damarukanadhan; M Polireddy; N Ravinder,"] Real Time Sentimental Analysis on Twitter", 2021 6th International Conference on Inventive Computation Technologies (ICICT).
- [11] Feng Liu;Wei-Wei Guo," Research on House Recommendation Model Based on Cosine Similarity in Deep Learning Mode in Grid Environment", 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS).
- [12] Kanta Nakamura; Kazushi Okamoto, "Directed Graph-based Researcher Recommendation by Random Walk with Restart and Cosine Similarity", 2020 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS).
- [13] Gurinder Singh;Bhawna Kumar; Loveleen Gaur;Akriti Tyagi," Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification",2019 International Conference on Automation, Computational and Technology Management (ICACTM).
- [14] Prabha PM Surya;Lakshmi V Seetha; B Subbulakshmi," Analysis of user emotions and opinion using Multinomial Naïve Bayes Classifier", 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA).
- [15] Noor Ifada;Triyani Fatchur Rahman;Mochammad Kautsar Sophan," Comparing Collaborative Filtering and Hybrid based Approaches for Movie Recommendation", 2020 6th Information Technology International Seminar (ITIS).
- [16] Rujhan Singla;Saamarth Gupta;Anirudh Gupta;Dinesh Kumar Vishwakarma,"] FLEX: A Content Based Movie Recommender", 2020 International Conference for Emerging Technology (INCET).

- [17] Gaurav Jariwala;Harshit Agarwal;Vrai Jadhav,” Sentimental Analysis of News Headlines for Stock Market”, 2020 IEEE International Conference for Innovation in Technology (INOCON).
- [18] P.P Gokul;B K Akhil;Kumar K. M Shiva,” Sentence similarity detection in Malayalam language using cosine similarity” 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).
- [19] Swe Swe Nyein,” Mining contents in Web page using cosine similarity” 2011 3rd International Conference on Computer Research and Development.
- [20] Md. Tayeb Himel;Mohammed Nazim Uddin;Mohammad Arif Hossain;Yeong Min Jang, “Weight based movie recommendation system using K-means algorithm” 2017 International Conference on Information and Communication Technology Convergence (ICTC).
- [21] Shweta Rana; Archana Singh ; “Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques“,2016 2nd International Conference on Next Generation Computing Technologies (NGCT).
- [22] G. Saranya;G. Geetha;Chakrapani. K;Meenakshi K;S. Karpagaselvi,” Sentiment analysis of healthcare Tweets using SVM Classifier,”2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS).
- [23] B. K. Bhavitha;Anisha P. Rodrigues;Niranjan N. Chiplunkar,” Comparative study of machine learning techniques in sentimental analysis”,2017 International Conference on Inventive Communication and Computational Technologies.
- [24] Vipin Kumar;Basant Subba,” A TfidfVectorizer and SVM based sentiment analysis framework for text data corpus”,2020 National Conference on Communications (NCC).
- [25] Raghad Obeidat;Rehab Duwairi;Ahmad Al-Aiad,” Collaborative Recommendation System for Online Courses Recommendations”, 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML).

- [26] M. Kanakaraj and R. M. R. Guddeti, "Performance analysis of Ensemble methods on Twitter sentiment analysis using NLP techniques", Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015).
- [27] Sajal Halder;A.M. Jehad Sarkar;Young-Koo Lee ,”Movie Recommendation System Based on Movie Swarm”, 2012 Second International Conference on Cloud and Green Computing.
- [28] Tianqi Zhou;Lina Chen;Jian Shen, “ Movie Recommendation System Employing the User-Based CF in Cloud Computing”, 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC).
- [29] Muyeed Ahmed;Mir Tahsin Imtiaz;Raiyan Khan,Movie Recommendation System using clustering and pattern recognition network”, 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC).
- [30] Sung Min Kim;Young Guk Ha,” A New Method of Measuring Document Similarity for Movie Recommendation”,2014 Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing.
- [31] R Lavanya;B. Bharathi,” Systematic analysis of Movie Recommendation System through Sentiment Analysis”, 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS).

CHAPTER 10

SELF ASSESSMENT OF PO-PSO ATTAINMENT

Table 6 :

Programme Outcomes (PO)	Justification
PO1. Engineering knowledge	We have been able to apply knowledge of science, mathematics and engineering like machine learning, statistics and analysis of the data for prediction, successfully.
PO2. Problem analysis	The prediction of the similar movies required to design machine learning models which best fit the data set. Experiments were to be performed on the models to compare and contrast the best model for the particular dataset.
PO3. Design/development of solutions	The analysis of the data set was carried on and then certain pre-processing techniques were used to refine the data before applying the machine learning models.
PO4. Conduct investigations of complex problems	The pre-processing techniques are used in certain different combinations as per the analysis of the data. The complex task at hand was a comprehension of the working of various pre-processing methods and machine learning models as we had to identify the models for our data.
PO5. Modern tool usage	The data are processed and analysed on

	Jupyter Notebooks which is a multi-language collaborative computing environment using Python as a programming language. Various tools like numpy, pandas, scikit etc are utilized.
PO6. The engineer and society	We have taken great care to maintain our social, professional and moral responsibilities since the day we have undertaken the project. The responsibility, which is placed upon us as individuals, towards society is the prime reason behind trying to enhance the entertainment industry, from the knowledge we acquired as engineers.
PO7. Environment and sustainability	The solution for the problem statement was designed after careful considerations of environmental impact and the necessity for sustainable development.
PO8. Ethics	Design and development, communication and working together was conducted ethically without any blemishes.
PO9. Individual and team work	We had broken down the design into individual modules and assigned them to each other with mutual consent. The modules were assembled together effectively, with teamwork.
PO10. Communication	There were no communication barriers as we could easily talk to each other as and when needed. We worked together to clarify doubts and help each other

	throughout.
PO11. Project management and finance	As each of us had an equal amount of work, we took responsibilities of our work and managed each of them with great care. The softwares used for our project is mainly open-source softwares. Thus, we did not have any financial dependencies.
PO12. Life-long learning	There is a plenty of need for engineering solutions in the entertainment industry as well as the other industries. We will continue to contribute to the society we live in, with the knowledge we gain. We will always try to engage ourselves in solving problem statements and learn lifelong.

Table 7 :

Programme Specific Outcomes (PSO)	Justification
PSO 1: Professional Skills	We applied machine learning knowledge to analyse the data set and recognize the type of model to be applied on the data set with certain pre-processing. The models required programming knowledge which was applied efficiently. Outcomes and results were analysed proficiently.
PSO 2: Problem Solving Skills	Pre-processing techniques and the machine learning models best suited for the data set were recognized

	<p>resourcefully. Since the methodology used for the design and development of the solution was agile methodology, choosing of the techniques and models became much easier. It also helped us in developing models with different combinations of pre-processing techniques in order to create a robust solution.</p>
<p>PSO 3: Ethics and Career Development</p>	<p>We worked cordially in a team crossing no ethical or moral barriers. The responsibility and ability to work in a project of another discipline benefitted us with empathy and comprehension of other disciplines. We improved our interpersonal skills and professional ethics which will be fruitful in our careers.</p>

APPENDIX 1 : PAPER PROOF

PAPER :

An Emoji-Based Movie Recommendation System using Sentimental Analysis and Machine Learning Techniques

Karthik G¹, Kishore Kumar AR², K V Dharma Teja³, Nagaraju P⁴, Mohan BA⁵, and Shiva Darshan SL⁶

Department of Computer Science and Engineering
{karthikgopal678¹, kishorekumar2254²,
imdharmar28³,nagarajupericherla123⁴}@gmail.com,
{mohan.ba⁵, shiva.darshan⁶}@nmit.ac.in

Abstract. Currently, movie recommendation systems are predominant intelligent systems that play a crucial role in providing selective information. Conventional approaches in recommending solutions are collaborative filtering and content-based filtering that has certain constraints when involving multiple domains. These existing approaches prerequisites user history and emotions to help the user find the desired. In order to address these issues, this paper proposes a Movie Recommendation System (MRS) using sentiment analysis and displays the movie comment review in one word along with emoji by analyzing reviews from the user. We have used various machine learning algorithms for recommendation as well as sentiment analysis. Further, comparative analysis was made to find the better classifier considering certain standard performance metrics. Experimental results showed that the proposed sentiment-based model illustrates better performance using the Naive Bayes model.

Keywords: Correlation, Machine learning algorithms, Recommending systems. Sentiment Analysis.

PROOF OF SUBMISSION :

7/20/2021

Nitte Meenakshi Institute of Technology Mail - AITA-2021 submission 129



Dr. Shiva Darshan <shiva.darshan@nmit.ac.in>

AITA-2021 submission 129

1 message

AITA-2021 <aita2021@easychair.org>
To: Shiva Darshan SI <shiva.darshan@nmit.ac.in>

Wed, Jun 30, 2021 at 11:52 PM

Dear authors,

We received your submission to AITA-2021 (International Conference on Artificial Intelligence: Theory and Applications):

Authors : Karthik G, Kishore Kumar Ar, Dharma Teja K V, Nagaraju P, Mohan Ba and Shiva Darshan SI
Title : An Emoji-Based Movie Recommendation System using Sentimental Analysis and Machine Learning Techniques
Number : 129

The submission was uploaded by Dr. Shiva Darshan SI
<shiva.darshan@nmit.ac.in>. You can access it via the AITA-2021
EasyChair Web page

<https://easychair.org/conferences/?conf=aita2021>

Thank you for submitting to AITA-2021.

Best regards,
EasyChair for AITA-2021.

APPENDIX 2 : PLAGIARISM CHECK REPORT

ORIGINALITY REPORT			
9%	5%	7%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS
PRIMARY SOURCES			
1	Gurinder Singh, Bhawna Kumar, Loveleen Gaur, Akriti Tyagi. "Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification", 2019 International Conference on Automation, Computational and Technology Management (ICACTM), 2019 Publication	1%	
2	ijaiem.org Internet Source	1%	
3	hdl.handle.net Internet Source	<1%	
4	ieeexplore.ieee.org Internet Source	<1%	
5	Feng Liu, Wei-Wei Guo. "Research on House Recommendation Model Based on Cosine Similarity in Deep Learning Mode in Grid Environment", 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), 2019 Publication	<1%	

6	Ching-Seh Mike Wu, Deepti Garg, Unnathi Bhandary. "Movie Recommendation System Using Collaborative Filtering", 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 2018 Publication	<1 %
7	leuten-leiden.com Internet Source	<1 %
8	www.ijitee.org Internet Source	<1 %
9	cs.bath.ac.uk Internet Source	<1 %
10	Kanta Nakamura, Kazushi Okamoto. "Directed Graph-based Researcher Recommendation by Random Walk with Restart and Cosine Similarity", 2020 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS), 2020 Publication	<1 %
11	Nimish Kapoor, Saurav Vishal, Krishnaveni K S. "Movie Recommendation System Using NLP Tools", 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020	<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography On

