A

PROJECT REPPORT

ON

# Movie Recommender System

Submitted in partial fulfilment of the requirements

of the degree of

## Bachelor of Engineering In Information Technology

by

Rohan Surve (57)

Abhishek Singh (55)

Sudhanshu Mohite (29)

Ajit Sargar(47)

Supervisor(s):

Prof. Aarti Abhyankar



**Department of Information Technology**

K.C. College of Engineering and Management Studies And

Research, Thane (E)

University of Mumbai

2022-23

## CERTIFICATE

This is to certify that the project entitled **“Movie Recommender System”** is a bonafide work of **Rohan Surve (57), Abhishek Singh (55), Sudhanshu Mohite (29),**

**Ajit Sargar (47)** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **“Bachelor of Engineering”** in

**“Information Technology”**.

Name and sign Name and sign

Supervisor/Guide Co Supervisor/Guide



Prof.Amarja Adgaonkar Dr.Vilas Nitnaware

Head of Department Principal

**Project Report Approval for T.E.**

This project report entitled ***Movie Recommender System*** by ***Mr. Rohan Surve, Mr. Abhishek Singh, Mr. Sudhanshu Mohite, Mr. Ajit Sargar*** is approved for the degree of Bachelor of Engineering in **Information Technology.**

Examiners

1.---------------------------------------------

2.---------------------------------------------

Date:

Place:

# DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Rohan Surve(57)

Abhishek Singh(55)

Sudhanshu Mohite(29)

` Ajit Sargar(47)

Date:

# DECLARATION

I declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Rohan Surve(57) Abhishek Singh(55)

Sudhanshu Mohite(29)

` Ajit Sargar(47)

Date:

# ACKNOWLEDGEMENT

We would like to express special thanks of gratitude to our guide **Mrs. Aarti Abhyankar**  as well as our Project Coordinator **Dr. Kiran Bhandari** gave us the golden opportunity to do this wonderful project on the topic of **Movie Recommender System**, which also helped us in doing a lot of research and we came to know about so many new things. We are very grateful to our Head of the Department **Mrs.Amarja Adgaonkar** for extending her help directly and indirectly through various channels in our project work**.** We would also like to thank Principal

**Dr. Vilas Nitnaware** for providing us the opportunity to implement our project. We are really thankful to them. Finally we would also like to thank our parents and friends who helped us a lot in finalizing this project within the limited time frame.

Thanking You.

## TABLE OF CONTENT

**Sr.No. Topic Page No.**

**Certificate ..............................................................................................i**

**Approval Sheet .....................................................................................ii**

**Declaration …….......................……………………………………….iii**

**Acknowledgement ……………………………………..…..................iv**

**List of Figures ...………………………………………………………vi**

**List of Table ...………………………………………………………...vii**

**Abstract ....…………………………………………………………….viii**

1. **Introduction …………………………………………………………………1**
2. **Literature Survey…………………………………………………………………..**
3. **Proposed Work** 
   1. Requirement Analysis ……………………………………………….
      1. Scope ……………………………………………….
      2. Feasibility Study ………………………………………
      3. Hardware & Software Requirement …………………………
   2. Problem Statement ………………………………………………………..
   3. Project Design……………………………………………………………...
   4. Methodology………………………………………………………………

**5. Conclusion and Future Scope…………………………………………………..**

**References ……………………………………………………………....**

## LIST OF FIGURES

**Sr.No. Topic Page No.**

1. Flow Chart 15

1. Block Diagram 15
2. DFD Level 0 16

1. DFD Level 1 17
2. Use Case Diagram 18

1. Output User Interface 17

# ABSTRACT

A recommendation engine filters the information mistreatment totally different algorithms and recommends the foremost relevant things to users. It 1st captures the past behaviour of a client and supported that, recommends product that the users can be seemingly to shop for. If a totally new user visits an e-commerce website, that website won't have any past history of that user. therefore however will the positioning approach advocating product to the user in such a scenario? One attainable answer might be to recommend the popular product, i.e. the product that are high in demand.

Another attainable answer might be to advocate the product which might bring the most profit to the business. 3 main approaches are used for our recommender systems. One is Demographic Filtering i.e they provide generalized recommendations to each user, supported picture show quality and/or genre. The System recommends identical movies to users with similar demographic options. Since every user is totally different, this approach is taken into account to be too straightforward.

The basic plan behind this technique is that movies that ar a lot of common and critically acclaimed can have the next likelihood of being likeable by the common audience. Second is content-based filtering, wherever we have a tendency to try and profile the user’s interests mistreatment data collected, and advocate things supported that profile. the opposite is cooperative filtering, wherever we have a tendency to try and cluster similar users along and use data regarding the cluster to create recommendations to the user.

## 1. INTRODUCTION

Recommender systems used in a various form of areas together with movies, music, news, books, analysis articles, search queries, social tags, and merchandise normally.

Recommendation System is a filtration program whose prime goal is to predict the movie to a user towards a domain-specific item. In our case, this domain-specific item is a movie, so the most focus of our recommendation system is to filter and predict solely those movies that a user would favor given some information concerning the user him or herself. There are many alternative ways that to create movie recommendation system however we've selected the content base recommender system in order that user will simply get the foremost similar movies on the user's interest. As our recommender system recommends the top high five movies as like movie that user is selected.

Earlier, the users needed to settle on choices on what books to purchase, what music to tune in to, what motion pictures to watch and so on. Commercial movie libraries effectively exceed 15 million films, which boundlessly exceeds the visual ability of any single individual. With a large number of motion pictures to browse, individuals now and then get overpowered. Therefore, an efficient recommendation system is necessary for the enthusiasm of both movie service providers and customers . With the improvement of recommendation systems, the customers will have no agony in settling on choices and organizations can keep up their client gathering and draw in new clients by improving users' satisfaction . Additionally, nowadays the modern technologies like machine learning and deep learning also plays a vital role in the process flexible technologies for day to day operations. In this manuscript, we discuss about the recommendation by using machine learning. Now, we discuss a method that has been previously implemented.

## 2. LITERATURE SURVEY

A movie recommendation system based on collaborative filtering approaches. Collaborative filtering takes the data from all the users and based on that generates recommendations. A hybrid system has been presented by Virk et al. [30]. This system combines both collaborative and content-based method. De Campos et al. [34] also made an analysis of both the traditional recommendation techniques. As both of these techniques have certain setbacks, he proposed another system which is a combination of Bayesian network and collaborative technique. a proposed clustering as an approach to handle the recommendations. Here a tree had been built which shows the user recommendation. Some of the major contribution in recommendation system is discussed in ***Table 1***.

***Table 1.*** Literature review of recommendation systems.

|  |  |  |
| --- | --- | --- |
| Authors | Year | Descriptions |
| Scharf & Alley [38] | 1993 | The authors proposed a flexible multicomponent rate recommendation system to predict the optimum rate of fertilizer for winter wheat. |
| Basu et al. [39] | 1998 | The authors proposed an approach to the recommendation that can exploit both ratings and content information. |
| Sarwar et al. [40] | 2001 | The authors proposed various techniques for computing itemitem similarities. |
| Bomhardt [41] | 2004 | The author proposed an approach for a personal recommendation of news. |
| Manikrao &  Prabhakar  [42] | 2005 | The authors presented the design of a dynamic web selection framework. |
| Von Reischach et al. [43] | 2009 | The authors proposed a rating concept that  allows users to generate rating criteria. |
| Choi et al. [44] | 2012 | The authors proposed approaches for integrating various techniques for improving the recommendation quality. |

***Table 2*** discussed the contribution of filtering techniques for different purposes.

***Table 2.*** Literature review of filtering techniques.

|  |  |  |
| --- | --- | --- |
| Authors | Year | Descriptions |
| Goldberg et al. [45] | 1992 | The authors introduced the collaborative filtering technique. |
| Herlocker et al. [46] | 1997 | Authors applied filtering techniques to Usenet news. |
| Miyahara & Pazzani [47] | 2000 | The authors introduced an approach to calculate the similarity between a user from negative ratings to positive ratings separately. |
| Hofmann [48] | 2004 | The author introduced a new-family of model-based algorithms. |
| Dabov et al. [49] | 2008 | The authors proposed an image restoration  technique using collaborative filtering. |
| Pennock et al. [50] | 2013 | The authors proposed various approaches for filtering by personality diagnosis. |
| Liu et al. [51] | 2014 | The authors introduced a new method to provide an accurate recommendation. |

One undertaking from the above discussions is that recommendations systems have gained vital name and recognition among researches because of their frequent look in varied and widespread applications within the fields of various branches of science and technology.

The previous recommendation systems had certain gaps in them:

* As most of the users do not provide ratings, the rating matrix becomes very sparse.
* Over-specialization is the most common problem faced by content-based recommendation.
* Content-based recommendation systems always face the problem of a cold start.

Therefore, this motivates us to provide a new model for society:

* Improves sparsity by making rating as mandatory.
* The problem of over-specialization is resolved using neighborhood-based collaborative techniques.

## 3. PROPOSED WORK

**3.1 Requirement Analysis:**

**3.1.1 Scope:**

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.

User Experience Enhancement:

The primary goal of a movie recommender system is to enhance the user experience by providing personalized movie recommendations based on their preferences, viewing history, and behavior patterns.

Content Discovery:

Recommender systems help users discover new movies they might enjoy by analyzing their past interactions, ratings, and feedback, as well as considering similarities between movies.

Platform Integration:

Movie recommender systems can be integrated into various platforms such as streaming services, e-commerce websites, mobile apps, and social media platforms to offer personalized recommendations to users.

Algorithmic Diversity:

The scope includes exploring and implementing different recommendation algorithms such as collaborative filtering, content-based filtering, matrix factorization, deep learning-based models, and hybrid approaches to improve recommendation accuracy and diversity.

**3.1.2 Feasibility Study:**

movie recommender system through machine learning involves several crucial considerations to determine its viability. Firstly, market analysis entails identifying the target audience and examining the current demand for such services along with potential competitors .From a technical perspective, the availability of pertinent data for model training, algorithm complexities, and resource requirements must be assessed. Libraries and frameworks conducive to building the recommender system should also be explored. Financial feasibility involves estimating initial investments, revenue streams, and conducting a cost-benefit analysis to ascertain the project's profitability and time to break-even. Legal and ethical dimensions necessitate adherence to data privacy laws, copyright regulations, and ethical concerns like bias mitigation and user privacy. Operational feasibility examines integration capabilities, scalability, and user experience, ensuring smooth functionality and satisfaction.

Technical Feasibility:

Evaluate whether the required technology and expertise are available to develop and maintain the recommender system. Assess the feasibility of implementing various recommendation algorithms and integrating them with existing platforms or systems.

Market Feasibility:

Analyze the target market for the movie recommender system, including demographics, user preferences, and existing competition. Determine whether there is a demand for personalized movie recommendations and if the system can gain traction among users.

Financial Feasibility:

Estimate the costs associated with developing, deploying, and maintaining the recommender system, including hardware, software, personnel, and ongoing operational expenses. Conduct a cost-benefit analysis to determine if the potential benefits outweigh the investment.

**3.1.3 Hardware & Software Requirement:**

A. Hardware Requirements B. Software Requirements

1. A PC with Windows/Linux OS 1. Anaconda distribution package
2. 2gb Graphic card (Jupyter notebook)
3. Processor with 1.7-2.4gHz speed 2. Anaconda distribution package
4. Minimum of 8gb RAM (PyCharm Editor)

3. Python libraries- Streamlit, Numpy,

. pandas, Requests

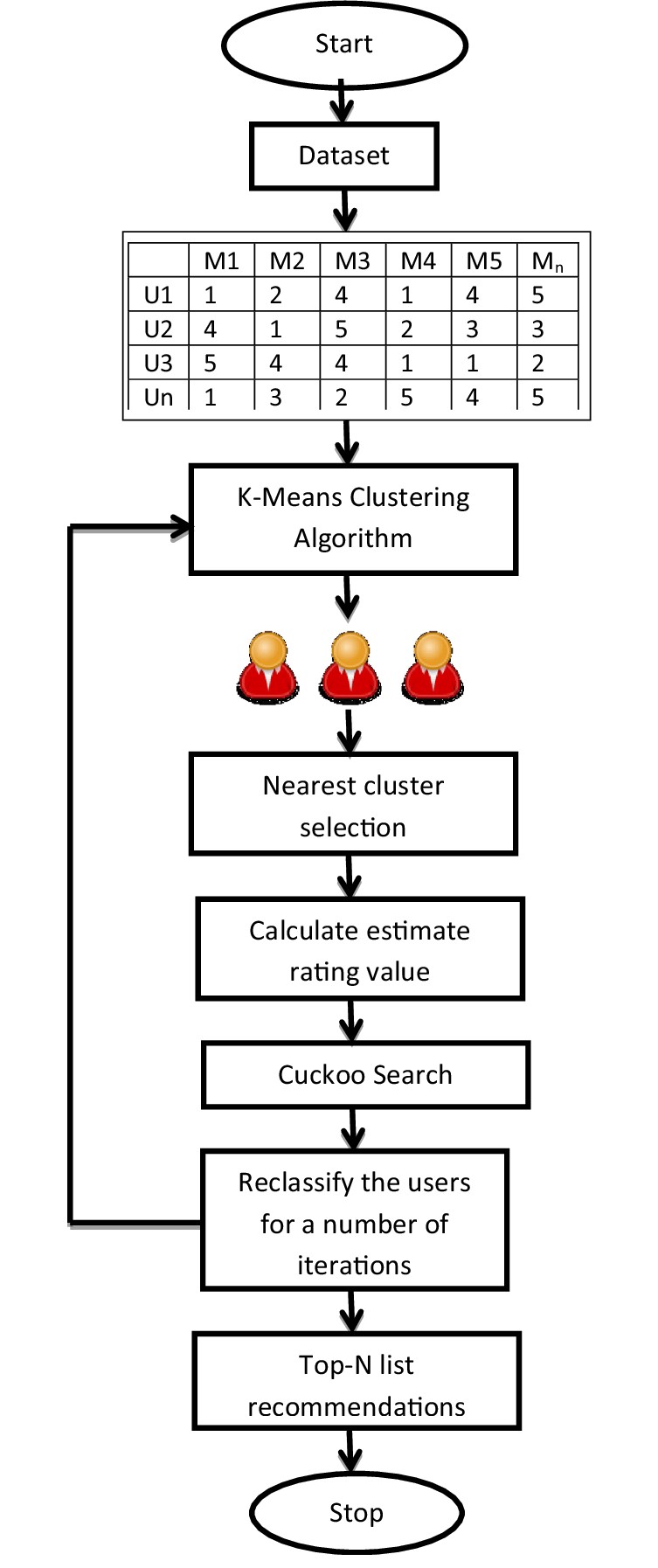
* 1. **PROBLEM STATEMENT**

This recommendation system recommends different movies to users. Since this system is based on a collaborative approach [67], it will give progressively explicit outcomes contrasted with different systems that are based on the content-based approach. Contentbased recommendation systems are constrained to people, these systems don't prescribe things out of the box. These systems work on individual users’ ratings, hence limiting your choice to explore more. While our system which is based on a collaborative approach computes the connection between different clients and relying upon their ratings, prescribes movies to others who have similar tastes, subsequently allowing users to explore more [68]. It is a web application that allows users to rate movies as well as recommends them appropriate movies based on other's ratings.

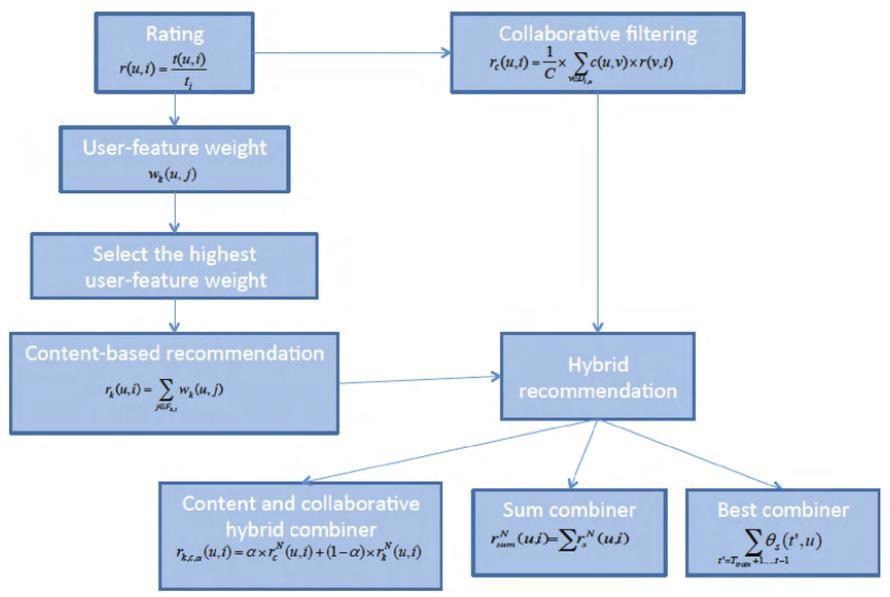
Despite the vast array of movies available across various platforms, users often struggle to discover content that aligns with their preferences and interests. Traditional browsing methods are inefficient and overwhelming, leading to user frustration and decreased engagement. There is a need for an intelligent movie recommender system that can accurately analyze user preferences and behavior to provide personalized recommendations, thereby enhancing user satisfaction and optimizing content discovery.

This problem statement succinctly outlines the primary issue (inefficient movie discovery), identifies the target audience (users seeking personalized recommendations), and highlights the desired outcome (enhanced user satisfaction and content discovery optimization). It sets the stage for the development of a movie recommender system by emphasizing the importance of addressing user needs and preferences through intelligent recommendation algorithms

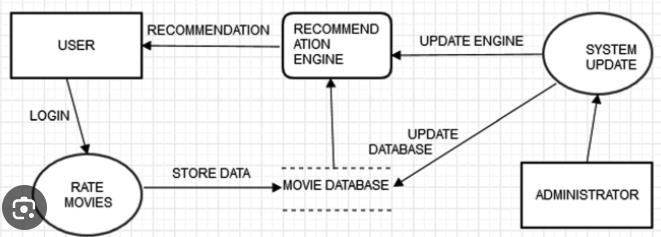
* 1. **PROJECT DESIGN**



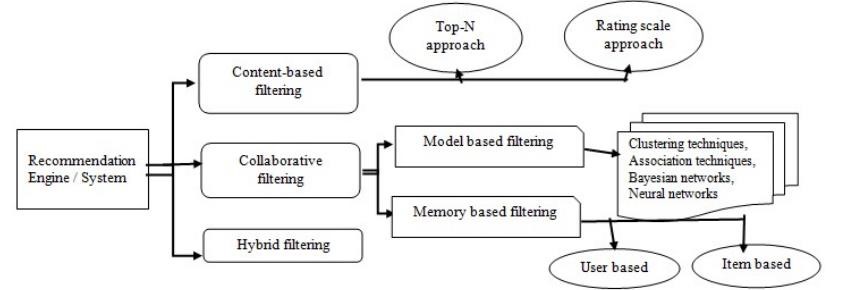
*Figure 1 : Flowchart*



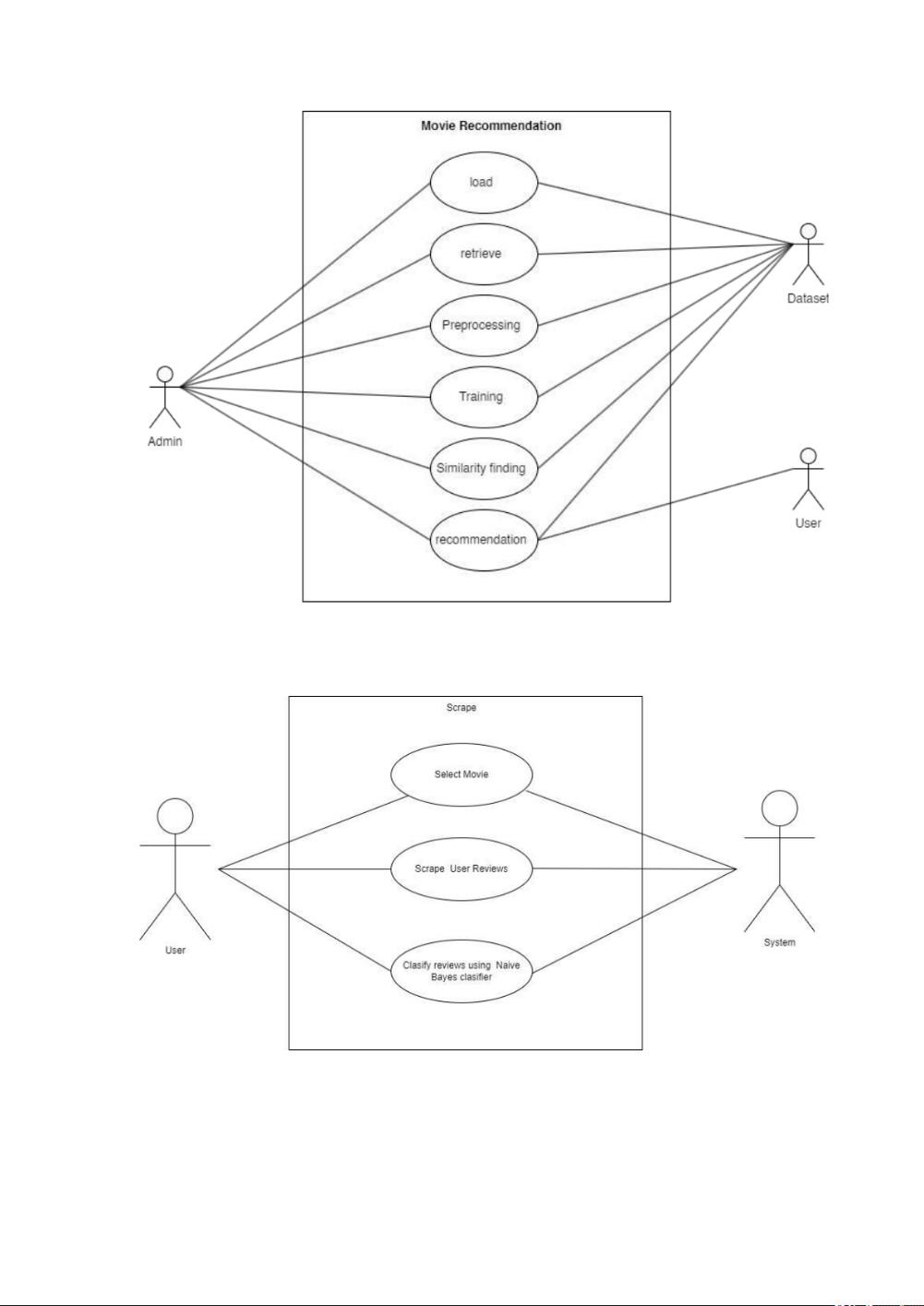
**Figure 2: Block Diagram**



***Figure 4: DFD Level 0***



**Figure 5: DFD Level 1**



*Figure 4 : Use Case Diagram*

### Code:-

**3.4 METHODOLOGY**

The methodology of the project is meant in six steps:

* Installing the Python and SciPy platform. we want to mount our ".ipynb" file on our google drive for more access.
* Loading the dataset. The dataset of picture show recommendation is required to be foreign in ".csv" format.
* Summarizing the dataset. Sorting and improvement of knowledge is that the necessary method to extend the potency of the project. we are able to fill the missing information victimisation "imputer" perform.
* Visualizing the dataset. we are able to visualize our "tmdb\_5000\_movies.csv" and "tmdb\_5000\_credits dataset through the Kaggle.com and so pre process method thereon.
* Evaluating some algorithms. when visualising the dataset, currently comes coaching and testing part!!! Let's divide the information the info the information) into 7:3 magnitude relation wherever seventieth data are trained and half-hour are tested. Now, let's choose the suitable models and so train them to urge the accuracy of the prediction. we've got used two models: COUNT VECTORIZER AND cos SIMILARITY. when obtaining the accuracy of every model and scrutiny them, lets cross Making some predictions. Now, comes the last stage of the project, i.e., to form predictions. Here, user will manually provide the input and acquire the advice of flick as per his/her interest.

* For content-based recommender system specifically, we have a tendency to conceive to notice a brand new thanks to improve the accuracy of the representative of the flick and suggest high 5 similar flicks to the user as per the interest of movie. Now, to form the project additional easy, we've got designed a frontend as well!!

The face consists of a web site with functions particularly recommendation and show. The face will be created victimisation flask module in python and preparation victimisation Heroku to link.

## 4. CONCLUSION & FUTURE SCOPE

**4.1 CONCLUSION:**

In this project, we have made the movie recommender system which recommends movies based on the content that is the most similar to what the user searches on the site. We have used cosine similarity to find out which top 5 movies would be the closest to what the user searches. This system tries to save up the time for the users who want to watch a movie similar to watch what they had watched before. We have used movie datasets from Kaggle.com which contained information like genre, cast, movie title, keywords in the movie, language and any more to preprocess the data and filter and gather all the information that would to be required to find out all the relevant data for content-based filtering.

4.2 FUTURE WORK:

In the proposed approach, we have used content based filtering, further we can move onto the hybrid filtering which would contain both content based and collaborative based filtering in the model. We can also put up and options for register/login so that the information of the users would be stored in the database and can be further used for collaborative filtering.

Integration of Advanced Recommendation Techniques:

Explore advanced recommendation techniques such as reinforcement learning, contextual bandits, or deep learning-based models to improve recommendation accuracy and relevance. Investigate how these techniques can capture complex user preferences and behavior patterns more effectively.

Dynamic and Real-time Recommendation:

Develop mechanisms for real-time recommendation updates based on user interactions and feedback. Implement algorithms that can adapt to changing user preferences and emerging trends in the movie industry to provide more timely and relevant recommendations.

Multi-modal Recommendation:

Incorporate multi-modal data sources such as images, audio, and text to enrich movie representations and enhance recommendation quality.

## 5. REFERENCES

1. Mohapatra, H., Panda, S., Rath, A., Edalatpanah, S., & Kumar, R. (2020). A tutorial on powershell pipeline and its loopholes. International journal of emerging trends in engineering research, 8(4), 975-982.
2. Kumar, R., Edalatpanah, S. A., Jha, S., & Singh, R. (2019). A Pythagorean fuzzy approach to the transportation problem. Complex & intelligent systems, 5(2), 255263.
3. Smarandache, F., & Broumi, S. (Eds.). (2019). Neutrosophic graph theory and algorithms. Engineering Science Reference.
4. Kumar, R., Edalatpanah, S. A., Jha, S., & Singh, R. (2019). A Pythagorean fuzzy approach to the transportation problem. Complex & intelligent systems, 5(2), 255263.
5. Mohapatra, H. (2009). HCR using neural network (Doctoral dissertation, Biju Patnaik University of Technology). Retrieved from

<https://www.academia.edu/39142624/HCR_USING_NEURAL_NETWORK>

1. Mohapatra, H., & Rath, A. K. (2019). Detection and avoidance of water loss through municipality taps in India by using smart taps and ICT. IET wireless sensor systems, 9(6), 447-457.
2. Mohapatra, H., & Rath, A. K. (2019). Fault tolerance in WSN through PE-LEACH protocol. IET wireless sensor systems, 9(6), 358-365.
3. Mohapatra, H., Debnath, S., & Rath, A. K. (2019). Energy management in wireless sensor network through EB-LEACH (No. 1192). Easy Chair.
4. Nirgude, V., Mahapatra, H., & Shivarkar, S. (2017). Face recognition system using principal component analysis & linear discriminant analysis method simultaneously with 3d morphable model and neural network BPNN method. Global journal of advanced engineering technologies and sciences, 4(1), 1-6.
5. Panda, M., Pradhan, P., Mohapatra, H., & Barpanda, N. K. (2019). Fault tolerant routing in Heterogeneous environment. International journal of scientific & technology research, 8(8), 10091013.
6. Mohapatra, H., Rath, A. K. (2020). Fault-tolerant mechanism for wireless sensor network. IET Wireless sensor systems, 10(1), 23-30.