

CERTIFICATE

Certified that **Aman Dhiman (2200290140023)**, **Aman Kumar (2200290140024)** have carried out the project work having “**FILMFLOCKS FINDER**” (**Project-KCA451**) for **Master of Computer Application** from Dr. A.P.J. Abdul Kalam Technical University (AKTU) (formerly UPTU), Lucknow under my supervision. The project report embodies original work, and studies are carried out by the student himself/herself and the contents of the project report do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Date: 27/05/2024

AMAN DHIMAN (2200290140023)

AMAN KUMAR (2200290140024)

.....
.....

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date: 24/02/2024

Ms. Divya Singhal
Assistant Professor
Department of Computer Applications
KIET Group of Institutions, Ghaziabad

Dr. Arun Tripathi
Head
Department of Computer Applications
KIET Group of Institutions, Ghaziabad

FLIMFLICKS FINDER
Aman Dhiman (2200290140023)
Aman Kumar (2200290140024)

ABSTRACT

This project report details the creation and execution of Film Flicks Finder, a sophisticated recommendation system tailored to offer personalised movie recommendations in the age of excessive information. The system uses collaborative and content-based filtering methods, incorporating machine learning algorithms to examine user behaviours, ratings, and viewing histories. The paper offers a thorough examination of the research framework, encompassing system architecture, data pre-processing, feature engineering techniques, and model selection and design.

Text processing techniques like stemming, bag-of-words and TF-IDF are used to analyse textual movie data. The system improves recommendation accuracy by evaluating film similarities through algorithms such as cosine similarity and Euclidean distance. The research concludes by proposing future directions that focus on exploring advanced machine learning techniques, integrating social media, enhancing content support, and refining the evaluation framework.

Film Flicks finder is an innovative recommendation system that provides personalised and accurate movie suggestions in the extensive and varied world of cinema. The study contributes to the continuous development of recommendation systems, focusing on the difficulties presented by the increasing content options and the requirement for personalised recommendations in the current rapidly changing digital landscape.

ACKNOWLEDGEMENTS

Success in life is never attained single-handedly. My deepest gratitude goes to my project supervisor, **Ms. Divya Singhal** for her guidance, help, and encouragement throughout my project work. Their enlightening ideas, comments, and suggestions.

Words are not enough to express my gratitude to Dr. Arun Kumar Tripathi, Professor and Head, Department of Computer Applications, for his insightful comments and administrative help on various occasions.

Fortunately, I have many understanding friends, who have helped me a lot on many critical conditions.

Finally, my sincere thanks go to my family members and all those who have directly and indirectly provided me with moral support and other kind of help. Without their support, completion of this work would not have been possible in time. They keep my life filled with enjoyment and happiness.

Aman Dhiman

Aman Kumar

.....

.....

TABLE OF CONTENTS

Certificate	i
Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	v
List of Figures	vi
1 Introduction	1-9
1.1 Overview	
1.2 The Dynamic Landscape of Recommendation Systems	
1.3 Application	
1.4 Challenges	
2 Comprehensive Work and Literature Review	10-17
2.1 Content-based filtering	
2.2 Collaborative filtering	
2.2.1 Memory based method	
2.2.2 Model based method	
3 Test Cases and Feasibility Study	18-21
4 Framework	22-45
4.1 Architecture	
4.2 Data preprocessing and feature engineering	
4.3 Exploratory Data Analysis (EDA)	
4.4 Text processing techniques	
4.5 Similarity Algorithms	
5 Implementations and Development	46-51
6 Discussion and Future Work	52-57
7 Conclusion	58-62
Bibliography	

LIST OF TABLES

Table No.	Name of Table	Page
1.2.1	Platform Specification Information	4
3.1	Test cases for movie recommendation system	21
4.2.1	Essential variables encompassed by the dataset	30

LIST OF FIGURES

Figure No.	Name of Figure	Page No.
1.4.1	Taxonomy of Movie Recommendation System	9
2.1.1	Content Based Filtering	12
2.2.1	Collaborative based Filtering	14
3.1	Test results for movie recommendation system	21
4.1.1	Flowchart of Proposed Methods	26
4.2.1	Information on Collected Data	29
4.3.1	Top 10 Genres by Revenue in the Dataset	32
4.3.2	Describe the quantitative columns	33
4.3.3	Comparison of Budget and Revenue (using Gathered Dataset)	33
4.3.4	Year Distribution (using the Gathered Dataset)	34
4.3.5	Distribution of Movies by Original Language (using the Gathered Dataset)	34
4.3.6	Runtime Distribution (using the Gathered Dataset)	35
4.3.7	Correlation Matrix of Quantitative Columns	35
4.3.8	Useful Data for Further Processing	36
4.3.9	Combined Multiple Features Columns into a Single Tags Column	37
4.3.10	Text Data of Tags Columns Converted into Small Letters	37
4.4.1	Porter Stemming Algorithm	38
4.4.2	Word Cloud of Stemmed Tag (using the Gathered Dataset)	39
4.4.3	Matrix of Tags Column After Applying the Bad-of Words Approach	39
4.5.1	Feature Matrix After Apply TF-IDF Approach	42
4.5.2	Prediction of Movies using Cosine Similarity Algorithm with Bow	43
4.5.3	Prediction of Movies using Cosine Similarity Algorithm with TF-IDF	44
4.5.4	Prediction of Movies using Euclidean Distance Algorithm with Bo	45
4.5.5	Prediction of Movies using Euclidean Distance Algorithm with TF-IDF	46
5.1	Home Page	50
5.2	Recommendation system	51
5.3	Recommended movies	51-52

