Movie Genre Classification Using Machine Learning

A.V. Siddharth[1], Pulipati Rakshitha[2], Shaik Mazher[3], V. Sumshaya Reddy [4]
Mrs. G. Uma Maheshwari[5]
[1,2,3,4]students, Department of CSE(AIML), Sreyas Institute of Engineering &Technology, Hyderabad, India- 500063
[5]Assistant Professor, Department of CSE(AIML), Sreyas Institute of Engineering &Technology, Hyderabad, India- 500063

***Abstract*—The film industry relies on genre classification of films to create recommendation systems and content organization with specific target audiences. Conventional classification methods rely on manual tagging that reflects subjective bias and inefficient operation. Automated and scalable movie classification by machine learning relies on text metadata like plot descriptions, reviews, and cast information. The system preprocesses movie descriptions via TF IDF vectorization before model training such as Logistic Regression, Naïve Bayes, Radom Forest, SVM. The system identifies the most precise model for prediction uses. The accuracy of 0.5836 was the highest for Logistic Regression when it classified 54,215 pre-specified IMDb datasets drawn from Kaggle- the proposed approach results in more accurate genre classification outcomes. Future studies should concentrate on applying deep learning techniques and larger datasets to enhance prediction accuracy. The system applies NLP methods such as Term Frequency Inverse Document Frequency (TF IDF), Word Embeddings and Transformer-based models like BERT to extract meaningful textual features. The suggested system obtains its data from publicly available movie descriptions sourced from IMDb and TMDb. This project applies automated genre classification using machine learning that provides increased accuracy and enhanced efficiency. The Proposed System employs machine learning to avoid manual tagging and uses a Multi-Model Approach to compare several classifiers rather than relying upon a single model and applying TF IDF vectorization for better text feature representation. The system contains Model Persistence functionality where trained models are saved for future application use. The system is more accurate since Logistic Regression and SVM work better than standard keyword-based classification.**

***Index Terms*—Logistic Regression, Naïve Bayes, Random Forest, SVM (Support Vector Machine), TF IDF vectorization.**

1.INTRODUCTION

As digital technologies continue to evolve, recommender systems have become increasingly important in helping users navigate a vast array of choices. These systems utilize machine learning algorithms trained on real-world data to detect patterns in user behavior and offer personalized suggestions. With the widespread use of the internet, especially social media, individuals now have platforms to share their opinions on news, events, entertainment, and more.

In particular, the film industry benefits significantly from these digital interactions. Instead of relying solely on traditional popularity metrics, insights are now drawn from blogs, review websites, and social platforms where both audiences and critics share their thoughts on movies. Understanding how viewers perceive a film is critical for production companies and investors, as it helps evaluate factors like storytelling effectiveness, scene impact, and overall reception. These insights can guide decisions about casting, marketing (such as trailer content), and even future productions, potentially leading to improved financial outcomes.

User reviews serve as a practical guide for prospective viewers trying to decide whether a film aligns with their interests. Platforms like IMDb (Internet Movie Database) are central to this process, offering extensive information on movies and TV shows, including cast details, ratings, genres, and viewer feedback. However, manually reviewing this vast amount of data is both labor-intensive and time-consuming. Machine learning and natural language processing (NLP) provide a solution by automating the analysis of these reviews to extract meaningful insights.

IMDb offers datasets that are widely used in research, particularly for sentiment analysis. One notable dataset, compiled by Maas et al. in 2011, includes 50,000 movie reviews labeled as positive or negative, and applies advanced NLP methods such as semantic similarity and probabilistic word models. Other datasets, such as Rounak Banik's "The Movie Dataset" (which focuses on metadata and ratings) and Lakshmi Pathi's "IMDb Dataset of 50K Movie Reviews" (which includes reviews and metadata), are also commonly used.

However, these datasets are not without issues. Since the information is web-scraped, it may lack consistency and fail to represent genres accurately. Reviews, often brief and opinionated, can be subjective and lead to class imbalance in training data. Mislabeling or biased data further complicates the effectiveness of sentiment analysis models. As a result, precision in class definitions and careful dataset curation are essential to ensure meaningful analysis.

II. LITERATURE SURVEY

1. Human tagging of humans manually.
2. Rule based or keyword based.
3. Prone to human inconsistencies and errors (Accuracy is lower)
4. Not easy to scale for big data sets (Scalability is lower)
5. Slow because it involves manual classification (Speed is lower)
6. 6 Limited to pre-defined rules (There is less flexibility)
7. No repetition; each new data set will need to be done manually.

Modern film genre classification systems leverage machine learning to analyze various types of data for accurate categorization. Text-based approaches utilize Natural Language Processing (NLP) techniques to interpret plot summaries, reviews, or scripts, often applying models like Support Vector Machines (SVM) or Long Short-Term Memory networks (LSTM) to predict genres. Image-based methods use Convolutional Neural Networks (CNNs) to analyze visual inputs, such as movie posters or scene frames, extracting features that are indicative of specific genres. Audio-based approaches rely on processing elements like dialogue or soundtracks using Mel-Frequency Cepstral Coefficients (MFCC) along with deep learning to identify genre-specific audio patterns. To enhance classification performance, many current systems integrate multiple data sources—text, visuals, and audio—into a multimodal framework. Additionally, ensemble techniques and algorithms like K-Nearest Neighbors (KNN) are commonly used to combine outputs from different models, thereby boosting accuracy. Platforms like IMDb and MovieLens implement these methods to organize films into genres and offer tailored viewing suggestions. However, challenges remain, particularly regarding ambiguous genre labels and the complexity of managing large-scale data.

Theory:

1.A Simplified Exploration Binary Relevance Method to Movie Genre Classification and Recommendations (Kumar et al.):

Think of this as training a classifier to recognize a movie's genre one by one. Each genre is assigned its own "yes or no" question, no matter how genres might relate to each other.

Although simple, this method might miss the subtlety of genre pairings, like "action comedy."

2.The Label Powerset Method:

Consider a movie as a blending of various genres, like mixing flavors. Each combination of genres is assigned a unique name according to this paradigm, which classifies "action comedy" and "drama horror" as entirely different genres.

3.Deep Learning and Multimodal Fusion (Ding et al., Behrouzi et al.):

Think of movies as three sensory experiences: sound (soundtrack), visuals (cinematography), and text (dialogues). Researchers create a system that "sees," "listens," and "reads" to correctly classify genres with neural networks such as RNNs.

4.Machine Learning Based Movie Suggestions (Ezeh et al., Appaji et al., Kaur et al.):

Through understanding your tastes, these algorithms seek to help you discover your next favorite movie. KNN, deep learning, and knowledge graphs' objective is to deliver customized recommendations depending on your watch history, cast, looks, and even genre of choice.

5.Movie Emotion Impact Prediction (Motamedi et al.):

Have you ever watched a movie that caused you to think about life or one that caused you to feel energized immediately? Hedonic scores focus on immediate pleasure, but eudaimonic scores measure that deeper satisfaction. By analyzing images, audio, and metadata, scientists can predict the emotions a film will evoke.

6.Generating Movie Scripts (Dharaniya et al.):
Making films from the ground up! These systems are capable of creating engaging dialogue and situations through the use of natural language processing and efficient deep learning. They write scripts by fusing creativity with algorithms.

7.Clustering and Classification Based on Reviews (González et al.):
Reviews contain a treasure trove of information. It is easier for researchers to classify films according to genre by grouping reviews; that is, they are able to use people's feedback to guide genre analysis.

8.Genre Based Recommender Systems (Vahed et al.):
It improves recommendations according to identified genres. It's like having a knowledgeable friend who always knows what you might like according to your previous interests.

Proposed System:

This model is designed to identify the most probable genre of a given movie by analyzing its plot summary. It automates the genre classification process using machine learning techniques. The system extracts key terms from the movie descriptions through the Term Frequency-Inverse Document Frequency (TF-IDF) method and processes this data through various machine learning models, including Logistic Regression, Naive Bayes, Random Forest, and Support Vector Machine (SVM), to determine which performs best. Each model is evaluated based on standard performance metrics such as accuracy, precision, recall, and F1-score. The model with the highest performance is selected for real-time genre prediction.

In experimental testing, Logistic Regression achieved the highest accuracy of 58.36% using a dataset of 54,215 movie summaries sourced from Kaggle, outperforming the other models. Although the overall performance is moderate, the results demonstrate the potential of this approach to enhance personalized movie recommendations. It is anticipated that incorporating a larger dataset and employing more advanced models, such as deep learning architectures, could lead to improved classification accuracy in future developments.

We used a number of machine learning models here :

1. Automation: it saves time on manual tagging by using machine learning.
2. Multi-Models: it compares multiple classifiers to pick the best one instead of depending on a single model.
3. Efficiency: TF IDF helps us to understand text better by converting the plots to a numerical feature.
4. Model Saving: Once trained, models are saved so we don't have to train them again.
5. Higher Accuracy: Logistics Regression and SVM gave better results than the traditional keywords-based system.

This will enhance the classification accuracy and efficiency. Making it suitable for real-world applications.

Algorithms:

1.Logistic Regression is a statistical method commonly used for solving binary classification tasks, where outcomes are typically in a "yes" or "no" format. It is widely used in applications such as spam filtering and medical diagnosis due to its simplicity and effectiveness in modeling binary outcomes.

2.Random Forest is an ensemble-based machine learning algorithm that builds multiple decision trees during training and combines their outputs to improve overall performance. It is highly effective for both classification and regression tasks, known for its robustness, ability to handle large datasets, and resistance to overfitting.

3.Support Vector Machine is a supervised learning algorithm used for both classification and regression challenges. It works by identifying the optimal hyperplane that best separates data points into distinct categories. The main objective of SVM is to maximize the margin between the decision boundary and the closest data points, improving the model’s generalization ability.

4.Naive Bayes is a simple yet powerful probabilistic classifier that works well with smaller datasets. It predicts movie genres based on the likelihood of word occurrences in plot descriptions. While it is fast and easy to implement, its performance is generally lower compared to more complex algorithms like Logistic Regression or Random Forest.

5.Neural Networks (Deep Learning)

- Feedforward Neural Networks are used for tasks such as genre classification by learning patterns in plot descriptions or movie ratings.
- Convolutional Neural Networks (CNNs) are more suitable for analyzing visual inputs like movie posters or scene frames, making them effective in visual-based classification tasks.
- Recurrent Neural Networks (RNNs) are ideal for handling sequential data, such as movie scripts or detailed plot summaries, due to their ability to capture temporal relationships.

These deep learning models are capable of recognizing complex patterns, especially in large-scale datasets. However, they require significant computational resources, large amounts of data, and are often more difficult to interpret compared to traditional machine learning models.

Differences:

- Complexity: More advanced models like Neural Networks and Random Forest can identify detailed patterns in data, in other hands low complex models like Logistics Regression and Naive Bayes are less simple and faster.

- Data Type: Such as, CNNs these are better for image data like movie poster while the RNNs and SVMs are good for text data like plots, movie description.

- Performance: Random Forest & Deep Learning models will give better results but it can be more computation power and time.

- Interpretability (Understanding the model):

Like Decision Tree and Logistics Regression model are easy to understand. When compared to complicated models like Neural Networks and Boosting methods are harder to explain and called "Black box" models.

Selecting the Right Model:

- By looking for larger datasets with complex patterns, the Random Forest algorithm might be suitable.
- for text data like movie description, review.We use SVMs and RNNs this would work fine.
- for image classification like poster or movie trailers we use CNN for this work.
- Using Naive Bayes if you need fast and quick simple results but it might not accurate for detailed tasks (complex tasks).

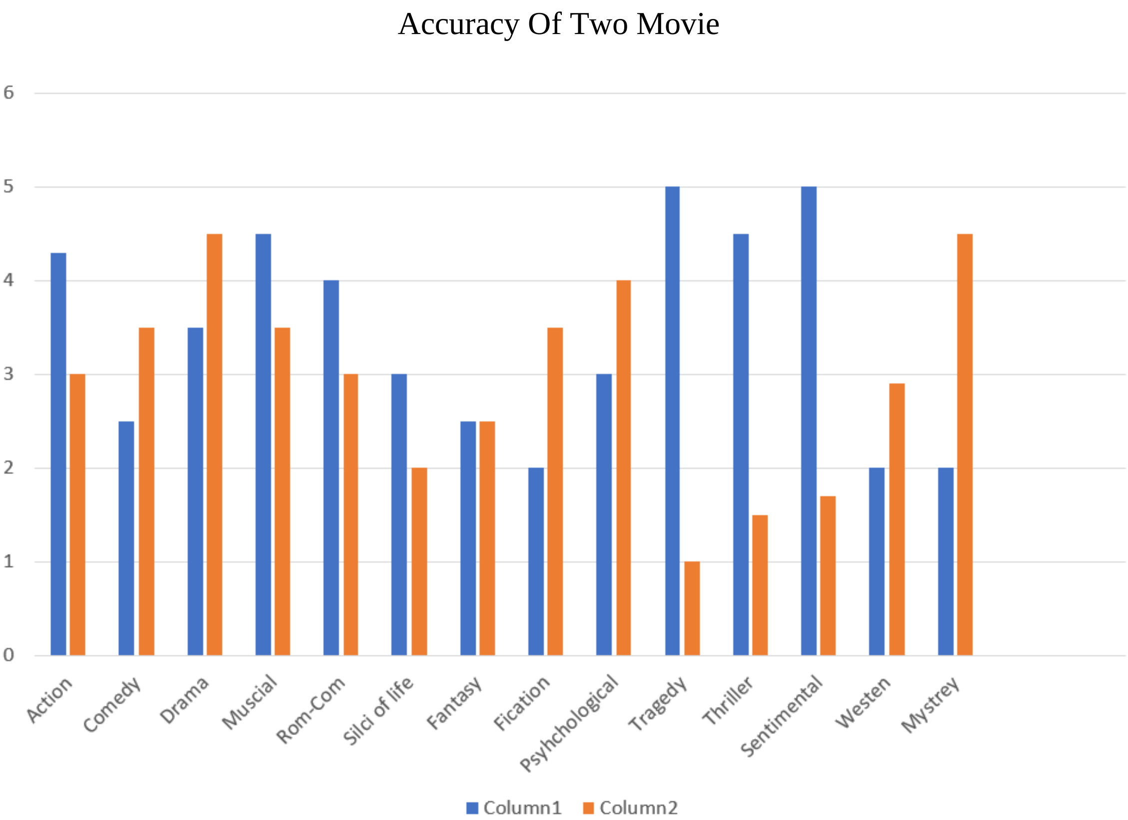
Every Model having its own strength and weaknesses. The right Model depends on how large and advanced your data is.

Performance of the algorithm is defined as it's accuracy, precision, recall and F1 score. That is moderated in this case. We find the Logistic Regression achieved as the great model from the high precision score that means it was the most accurate when classifying the genre for the movie

Table:

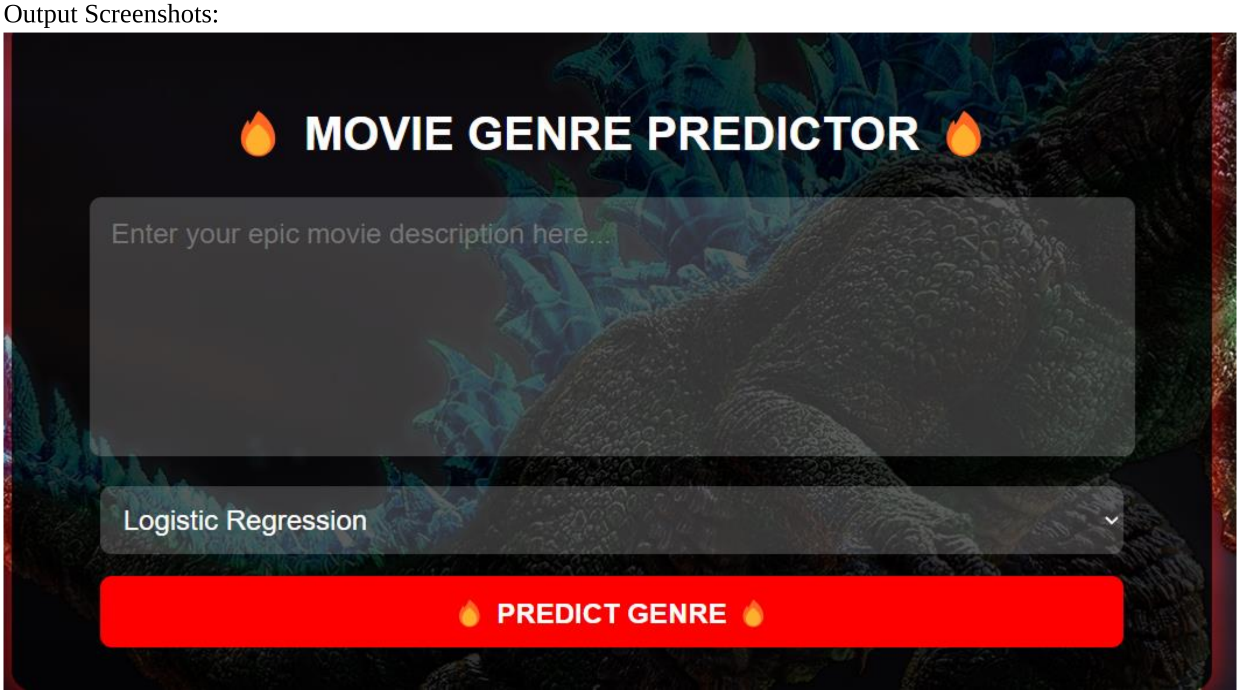
Model	Precision	F1 score	Accuracy
Logistic Regression	0.81	0.84	0.8480
Linear SVM	0.84	0.88	0.8847
Naïve Bayes	0.80	0.83	0.838

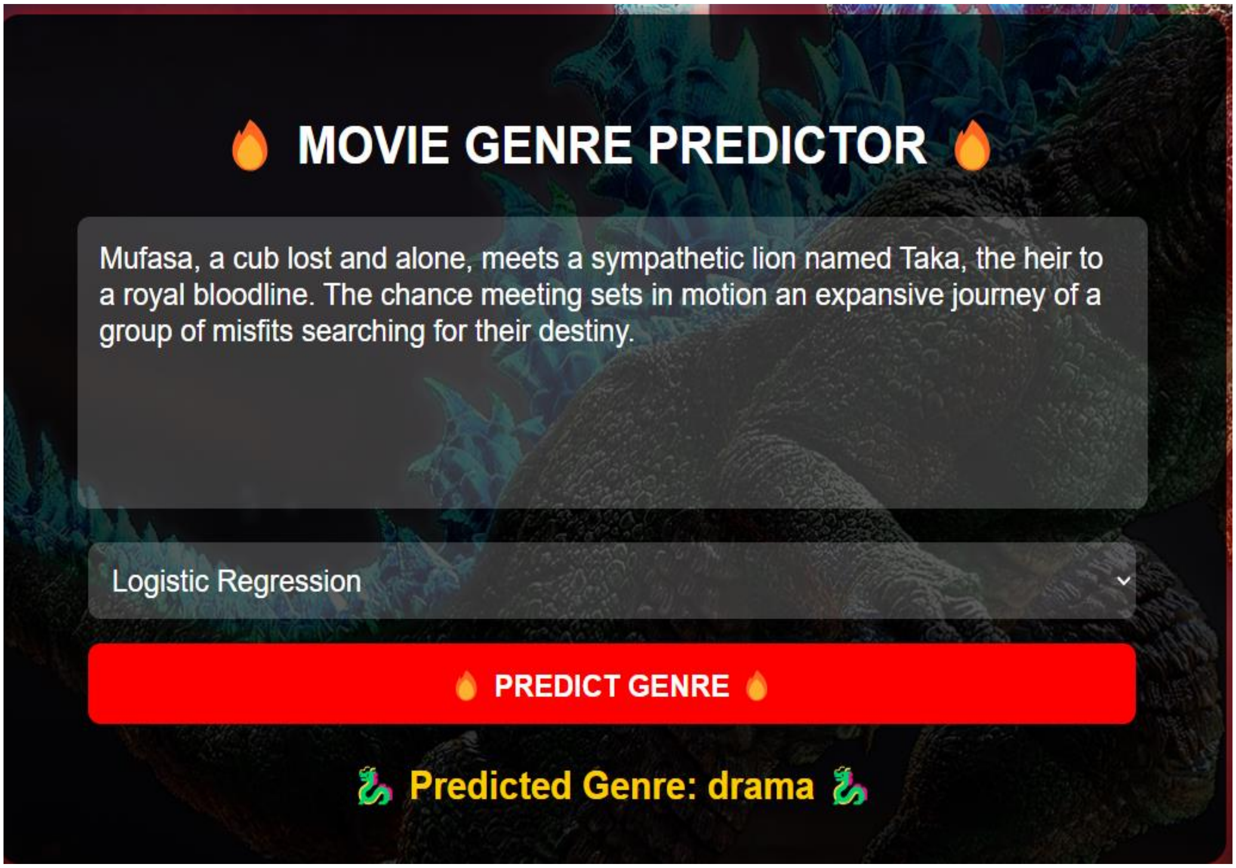
Graph:



Column1: Pusha2 Movie
Column2: MAD Movie

Output Screenshots:





III. CONCLUSION

In This project, we developed a system that identify the movie genre by using the information like reviews, summarize, caste details and description (ex: Action, Fantasy, love, etc.).

Instead of doing a manual tagging (this takes longer and depending on who is doing this).

--We are using the Machine Learning because it makes it quicker and consistent.

We use TF IDF (Term Frequency Inverse Document Frequency) to convert the text into numerical form that the machine can understand it. After this we test different machine learning model including the logistic Regression, Naive Bayes and SVM to see which will give the great results.

In this we select the Logistic Regression model, it achieved the accuracy of 58% on the dataset over 54,000 movies entries sourced from IMDb via kaggle. This consider as a best result by looking the difficulty of classifying the pattern text data for movie classification. By Using the Machine Learning model for movie genre, it has clearly advantage that it saves the time, reduce the mistakes that human might done it when categorizing the content. This is due to the machine learning algorithm, once it trained can reliably implement same criteria for classification in other words it keep things consistent.

Although our results are good. There still a room for improvements. The future versions of this system that could more advance tools like for example deep learning models "such as BERT" to understand the pattern behind the text even in a better way. BERT model developed to pick up the deeper patterns and context, that results in better accuracy.

We can improve results by implementing more dataset from multiple sources and by using individual techniques to represent the word in smarter way like the embedding. These all-change help make the system more accurate and prediction.

In short, the project shows that the machine learning models have a big role in sorting the movies by its genre in a smart, automatic way.it is a step towards improving how the content is recommended, managed and experienced by the user or viewers.

Author et al	Year of Publication	Algorithm Used	Implementation Details	Evaluation Parameter	Comments
Sanjay Kumar [1]	2023	Binary Relevance(BR), lable powerset(LP); Count Vectorizer naïve Bayes(MNB),K-Nearest Neighbor(KKN), Support Vector Classifier(SVC)	Multi-lable text classification for movie genre prediction using IMDb data; applied problem transformation techniques and text vectorization; evaluted 16 combinations with K-fold cross-validation	Accuracy,F1-score	Best performance:LP+TF-IDF+SVC with 0.95 accuracy and 0.86 F1-score
Motamedi Elham [2]	2024	Random Forest, Majority Classifier	Collected Eudaimonic and Hedonic scores for 709 movies(3699 records); augmented data with metadata, audio, low-level & high-level visual features; trained ML models	ROC AUC	Random Forest achieved 20% increase in ROC AUC over baseline;metadata features contributed most;automated scoring is a viable alternative to questionnaires
Ezeh Anthony [3]	2023	K-Nearest Neighbors	Developed a movie genre recommender system using TMDB 5000 Moive Dataset; integrated with a Streamlit web app allowing users to select moives and adjust similarity using a slider bar; varied K for flexibility	Accuracy,Precision,Recall,F1-score	Moderate to high precision; recall improves with K(max at K=15); average F1-score:0.600; personalized suggestions via interactive UI; useful for marketing and distribution insights
R. Dharaniya [4]	2023	TF-IDF,Word2Vec, Deep Belief Network,Bi-LSTM,GPT-	Proposed an Ensemble-based Moive Script Generation model using deep features	Standard Performance measures	Novel framework combinig feature engineering and advanced generative models;

		3,GPT Neo X,AI-CMO Optimizer	extracted from movie text data; ensemble of Bi-LSTM,GPT-3,GPT Neo X optimized by AI-CMO		aims to automate complex script generation using deep ensemble learning
S.Appaji [5]	2023	Content-Based Filtering, Collaborative Filtering, Hybrid Filtering	Study of popular moive recommender systems; analyzed filtering techniques based on user interaction, preferences, and item characteristics;reviewed similarity metrics and performance indicators	Not explicitiy stated	Focused on conceptual understanding and comparison of recooomendation strategies;highlights role of personalization in modern systems
Ihui & Hongwei Ding [6]	2024	Deep Learning	Proposed a multimodal genre classification framework using audio,poster,plot and video frame sequences;applied fusion via concatenation and element-wise sum;trained on LMTD-9 dataset	AU(PRC),AU(PRC)w	Outperformed state of the art by 8.6% and 5.3%; demonstrates the power of combining multiple data modalities for better classification accuracy
Gonzalez & Fernonado [7]	2023	Threshold Algorithm, Mini-Batch K-Means, TF-IDF	Proposed a threshold-based stopword removal method; used word frequency and presence conditions; applied mini-batch K-means on IMDb reviews with varied preprocessing; used TF-IDF for text representation	Mapping Genre Success Rate,Accuracy	Achieved80%+mapping success vs.Zipfs law method;final dataset accuracy improved to 94%;reclassification improves coherence and balance of sparsely labeled data
Nikhil Kumar	2022	Naïve Bayes	Pre-processed text from movie reviews; selected 500-20,000	Accuracy	Multinomaial Naïve Bayes perforamed best

Rajput [8]			words based on frequency using team frequency and TF-IDF for vector construction; applied 4 Naïve Bayes variants		with 86.46% accuracy;shows effectiveness of frequency-based word selection and TF-IDF in sentiment classification
Vahed & Pegah [9]	2024	Logistic Regression,Decision Tree,Navie Nayes,CNN+LSTM	Used movie dataset with titles,synopses,actors, genre lables;implemented in python using word2vec and TF-IDF for vectorization.	Fi-score	Achieved high genre prediction accuracy;comparable to state of the art;integrated web API for recommendations based on input title
Jyoti Tripathi & Sunita Tiwar [10]	2023	Multiple supervised ML algorithms	Features used: title, director, star cast, writer; Data preprocessed from IMDb	MSE, RMSE, R^2score, Accuracy=88.47%	Models compared; best-performing model chosen for integration based on performance metrices

REFERENCES

[1] Tripathi, Jyoti, Sunita Tiwari, Anu Saini, and Sunita Kumari. "Prediction of movie success based on machine learning and twitter sentiment analysis using internet movie database data." Indones. J. Electr. Eng. Comput. Sci 29 (2023): 1750-1757. d1wqtxts1xzle7.cloudfront.net/104511014/29246_60552_1_PB-libre.pdf?1690283074=&response-content-disposition=inline%3B+filename%3DPrediction_of_movie_success_based_on_mac.pdf&Expires=1744294629&Signature=Lp82kI64gtXz3h7NJrFEIOJA9aYXm--KEoWLFBznY~f-M7AbomKt9F8tPraGJJmVt7e8XDmxTGHU9tWTB~XNJum8A06e-0luJg65VBZIjMhyfOcHSEs2HWZeqBamQhm0TF5vtRH~x6~LpNW89IpAevPjsXGSjlSF84kPbSfNnkHgAlYzjoH0tjPcWu2ZoOn9jUklBe1FEEX3p1EqKrLKx3b3~Nl~4eW2wa4~WawgorSzZuUq-AWFs1l3lH5~D80ULQzW9uSPJmbOYNxX1xoXZYr0U-0GpxOrfGSRIwPmwTSkYZJHoViedSXRO00fFUmxScFshEjjyWbzuyjIPZO8~g__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

[2] Kumar, Sanjay, Nikhil Kumar, Aditya Dev, and Siraz Naorem. "Movie genre classification using binary relevance, label powerset, and machine learning classifiers." Multimedia Tools and Applications 82, no. 1 (2023): 945-968.

[3] Motamedi, Elham, Danial Khosh Kholgh, Sorush Saghari, Mehdi Elahi, Francesco Barile, and Marko Tkalcic. "Predicting movies' eudaimonic and hedonic scores: A machine learning approach using metadata, audio and visual features." Information Processing & Management 61, no. 2 (2024): 103610.

[4] Ezeh, Anthony. "Developing Machine Learning-based Recommender System on Movie Genres Using KNN." (2023).

[5] Dharaniya, R., J. Indumathi, and V. Kaliraj. "A design of movie script generation based on natural language processing by optimized ensemble deep

learning with heuristic algorithm." Data & Knowledge Engineering 146 (2023): 102150.

[6] Kaur, Gurinder, Fei Liu, and Yi-Ping Phoebe Chen. "A deep learning knowledge graph neural network for recommender systems." Machine Learning with Applications 14 (2023): 100507. (gnn)

[7] Appaji, S., M. Kusuma Patnaik, M. Kumar, K. Manoj Kumar, and K. Satish. "Movie recommendation system using machine learning." In AIP Conference Proceedings, vol. 2794, no. 1. AIP Publishing, 2023.

[8] ihui, Hongwei Ding, Jinlu Wu, Ying Xi, Xuemeng Wu, and Xiaohui Cui. "Multi-label movie genre classification based on multimodal fusion." Multimedia Tools and Applications 83, no. 12 (2024): 36823-36840.

[9] González, Fernando, Miguel Torres-Ruiz, Guadalupe Rivera-Torruco, Liliana Chonona-Hernández, and Rolando Quintero. "A Natural-Language-Processing-Based Method for the Clustering and Analysis of Movie Reviews and Classification by Genre." Mathematics 11, no. 23 (2023): 4735.

[10] Behrouzi, Tina, Ramin Toosi, and Mohammad Ali Akhaee. "Multimodal movie genre classification using recurrent neural network." Multimedia Tools and Applications 82, no. 4 (2023): 5763-5784.

[11] Vahed, Pegah, Azadeh Tabatabaei, and Leila Taherkhani. "A Model for Movie Classification and a Genre-based Recommender System." In 2024 10th International Conference on Web Research (ICWR), pp. 216-225. IEEE, 2024.

[12] Jassim, Mustafa Abdalrassual, Dhafar Hamed Abd, and Mohamed Nazih Omri. "Machine learning-based opinion extraction approach from movie reviews for sentiment analysis." Multimedia Tools and Applications (2024): 1-26.

[13] Salman, Ahmed Hussein, and Waleed Al-Jawher. "A New Multi-class Classification Method Based on Machine Learning to Document Classification." In 2023 16th International Conference on Developments in eSystems Engineering (DeSE), pp. 605-610. IEEE, 2023.

[14] Mangolin, Rafael B., Rodolfo M. Pereira, Alceu S. Britto Jr, Carlos N. Silla Jr, Valéria D. Feltrim, Diego Bertolini, and Yandre MG Costa. "A multimodal approach for multi-label movie genre classification." Multimedia Tools and Applications 81, no. 14 (2022): 19071-19096.

[15] Zhang, Zhongping, Yiwen Gu, Bryan A. Plummer, Xin Miao, Jiayi Liu, and Huayan Wang. "Movie genre classification by language augmentation and shot sampling." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 7275-7285. 2024.

[16] Behrouzi, Tina, Ramin Toosi, and Mohammad Ali Akhaee. "Multimodal movie genre classification using recurrent neural network." Multimedia Tools and Applications 82, no. 4 (2023): 5763-5784.

[17] Rajput, Nikhil Kumar, and Bhavya Ahuja Grover. "A multi-label movie genre classification scheme based on the movie's subtitles." Multimedia Tools and Applications 81, no. 22 (2022): 32469-32490.

[18] Oyewola, David Opeoluwa, and Emmanuel Gbenga Dada. "Machine learning methods for predicting the popularity of movies." Journal of Artificial Intelligence and Systems 4, no. 1 (2022): 65-82.

[19] Unal, Fatima Zehra, Mehmet Serdar Guzel, Erkan Bostanci, Koray Acici, and Tunc Asuroglu. "Multilabel genre prediction using deep-learning frameworks." Applied Sciences 13, no. 15 (2023): 8665.

[20] Cai, Zihui, Hongwei Ding, Jinlu Wu, Ying Xi, Xuemeng Wu, and Xiaohui Cui. "Multi-label movie genre classification based on multimodal fusion." Multimedia Tools and Applications 83, no. 12 (2024): 36823-36840.

[21] Unal, Fatima Zehra, Mehmet Serdar Guzel, Erkan Bostanci, Koray Acici, and Tunc Asuroglu. "Multilabel genre prediction using deep-learning frameworks." Applied Sciences 13, no. 15 (2023): 8665.

[22] Kordabadi, Mojtaba, Amin Nazari, and Muharram Mansoorizadeh. "A movie recommender system based on topic modeling using machine learning methods." International Journal of Web Research 5, no. 2 (2022): 19-28.

[23] Paulino, Marco Aurélio D., Yandre MG Costa, and Valéria D. Feltrim. "Evaluating multimodal strategies for multi-label movie genre classification." In 2022 29th International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 1-4. IEEE, 2022.

[24] Paulino, Marco Aurélio D., Yandre MG Costa, and Valéria D. Feltrim. "Evaluating multimodal strategies for multi-label movie genre classification." In 2022 29th International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 1-4. IEEE, 2022.
[25] Popat, Anannya, Lakshya Gupta, Gowri Namratha Meedinti, and Boominathan Perumal. "Movie poster classification using federated learning." Procedia Computer Science 218 (2023): 2007-2017.
[26] Gourammolla, Samheeta, and S. Gokila. "HCB Machine Learning Approach for Movie Recommendation System." In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1186-1190. IEEE, 2022.
[27] Akyon, Fatih Cagatay, and Alptekin Temizel. "Deep architectures for content moderation and movie content rating." arXiv preprint arXiv:2212.04533 (2022).
[28] Akbar, Jihadul, Ema Utami, and Ainul Yaqin. "Multi-label classification of film genres based on synopsis using support vector machine, logistic regression and naïve bayes algorithms." In 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), pp. 250-255. IEEE, 2022.
[29] Liu, Z., Wang, X. and Zhu, H., 2022, March. A new machine learning algorithm for users' movie recommendation. In CIBDA 2022; 3rd International Conference on Computer Information and Big Data Applications (pp. 1-4). VDE.
[30] Velingkar, Gaurang, Rakshita Varadarajan, and Sidharth Lanka. "Movie box-office success prediction using machine learning." In 2022 second international conference on power, control and computing technologies (ICPC2T), pp. 1-6. IEEE, 2022.