

Explainable Neural Network analysis on Movie Success Prediction

S Bhavesh Kumar¹ and Sagar Dhanraj Pande^{2,*}

^{1,2}School of Computer Science and Engineering, VIT-AP University Amaravati, Andhra Pradesh, India

Abstract

These days movies are one of the most important part of entertainment industry and back in the days you could see everyday people standing outside theatres, or watching movies in OTT platforms. But due to busy schedules not many people are watching every movie. They go over the internet and search for top rated movies and go to theatres. And creating a successful movie is no easy job. Thus, this study helps movie producers to consider what are the important factors that influence a movie to be successful. this study applied neural network model to the IMDb dataset and then due to its complex nature in order to achieve the local explainability and global explainability for the enhanced analysis, study have used SHAP (Shapley additive explanations) to analysis.

Keywords: SHAP, ANN, LIME, Glass box model, Black Box model, Explanations (global and local)

Received on 09 October 2023 accepted on 14 November 2023, published on 21 November 2023

Copyright © 2023 S. Bhavesh Kumar *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetsis.4435

*Corresponding author. Email: sagarpande30@gmail.com

1. Introduction

The movie industry is one of the biggest dynamic and constantly evolving parts of the global entertainment industry. Movie production involves a complex process that requires significant investment, planning, and creativity. After a lot of generational changes new technologies and various kinds of stories came into action and also, mankind moved from theatre stage plays to 4D movies and have come a long way. And producing any movie these days takes a lot of time and money and if all the movie goes and becomes a flop it would be a huge waste of resources. And time is one of the most precious resource for the people. With that in mind this many data analytics and researches have attempted to develop models to help them predict the success of a movie beforehand. They collect data from various sources such as social media, movie reviews, box office sales to analyse and predict the outcome of a movie.

This study explores the use of neural networks and applies the concept of deep neural network on the data acquired. But the complex mechanisms and calculations of deep

neural networks are very hard when the model gets complex. But if the model is simple it won't make accurate predictions enough. So, researches in the past stick to the model's being complex as it was also giving them results. But for a prediction, we need to have a reason why this prediction happens, which feature influenced the prediction or else everything is going to be a black box and sooner or later there will be a time where we can't recognise the problem that why is this model is not giving accurate predictions. Here comes the concept of SHAP (Shapley Additive explanations) into the action. With the SHAP we can calculate local explanations and global explanations and understand the working of the model and hence converting our black box to a glass box model. This study uses SHAP to explain the working of the neural network and extract the important features which influence the output the most.

The sections of this study are as follows: Discussed related work in movie success prediction and explainable AI in Section 2, the literature survey. The methodologies associated with the explainable AI, LIME and SHAP, as well as neural networks and the proposed approach, are discussed in Section 3. The discussions that are based on

the proposed model's results are discussed in Section 4. The model's future aspects and conclusions serve as the foundation for Section 5.

2. Literature Survey

Zhang and Skeina worked on using news analysis to predict movies in 2009 [1]. It was discovered that using news data produced results that were comparable to using data from IMDb. Far superior outcomes were gotten by consolidating IMDb and news information. In 2015, Lash and Zhao suggested a strategy for predicting movie decisions for investment [2]. Early decisions regarding investments in the film industry were aided by this work. Data from the past was used in this study. Matching who to what and when to what was one of this work's hallmarks. Box office sales were the main factor used to determine the profit. Then other than that around 2019 You-Jin Kim and group used the plot summaries which are made before starting of the shoot of the movie and used ELMO embedding to predict the movie success [3]. Based on the written summary of the movie's plot by the director, they proposed a deep learning-based method that uses the ELMO embedding technique and sentiment analysis results of sentences to predict a movie's success. In all out they fabricated three profound learning models an ELMO implanting and two consolidated profound learning models (a blended 1D CNN organization and a combined leftover bidirectional LSTM organization. Wikipedia provided the textual summary that they obtained. For predicting "successful," they received a F1 score of 0.68, while for predicting "unsuccessful," they received a score of 0.70. In 2022 Sandipan Sahu and group published a paper which used K-fold hybrid deep ensemble learning model to predict success of a movie in early-production stage [4]. They have collected the past 30 years data regarding of Indian movie information, especially all regional wise movies and their proposed model delivered 96% accuracy. Another article by Yuan Ni and colleagues using multi-model ensembles was published in 2022 as well [5]. Their final model outperformed XGBoost, LightGBM, and other models after being trained and tested on the dataset of the top 2000 movies. Wei Lu and his team released an IFOA-GRNN-based movie box office forecasting model in August 2022. Using an improved fruit fly optimisation (IFOA) generalised regression neural network (IFOA) model, they compared their findings to those obtained using FOA-GRNN, GRNN, KNN, and other models. The acquired root mean square mistake on anticipating film industry is 0.3412 and precision is around 90% [6]. In 2022, Gaurang V, Rakshita V and Sidharth Lanka have also published a paper based on Movie Box-Office Success Prediction [7]. Their model has capability of taking planning as a genre and they have developed their model using Random Forest Regression model, also providing recommended budget, runtime.

2.1 Explainable AI

Many significant methods have been implemented related to movie success prediction now also have also saw few papers regarding explainable AI on which our study and model is mainly dependent on. In 2021 Manze Guo and group published a paper regarding "Older Pedestrian Traffic Crashes Severity Analysis Based on an Emerging Machine Learning XGBoost" [8] in which they used an XGboost model on their dataset which had three types of accidents making it a multi class classification model and performed SHAP analysis. And in 2023 Sagar Pande and Aditya Khamparia have published a paper on "EXPLAINABLE DEEP NEURAL NETWORK BASED ANALYSIS ON INTRUSION DETECTION SYSTEMS" [9]. They put forth a theory that combines a deep neural network-aided conventional intrusion detection system with the interpretability of the model predictions. To further improve and optimise their model, they utilised SHAP's local explainability and global explainability as well as the well-known dataset NSL-KDD. In November 2019 Karim E.M., Ben Peachey Higdon, Ayşe Başar have tried interpreting financial time series with the SHAP values, their work showed that SHAP values can be useful in improving the prediction of accuracy as they naturally cluster datapoints according to feature importance[10]. In 2020 Liat Antwarg, Ronnie Mindlin Miller, Bracha Shapira, Lior Rokach applied SHAP to explain anomalies detected by an autoencoder an unsupervised model. Their proposed method extracts and visually depicts both the features that most contributed to the anomaly and those that offset the anomaly[11]. In 2020 Kary Främling wrote an article on Explainable AI which was applied on Autonomous agents and Multi agent Systems. They concluded that Explainable AI (XAI) has been a core research topic in AI for decades, and CIU, based on Decision Theory, extends the notions of importance and utility for non-linear models of AI systems, providing a universal and model-agnostic foundation for XAI [12]. In 2019 Alberto Fernandez and group wrote an article which shows Explainable AI on Fuzzy Systems. Their paper points out why evolutionary fuzzy systems are important from an explainable point of view when they began and why do we need them and what are they used for [13]. In 2018 Andreas Holzinger wrote an article "From Machine learning to Explainable AI" which showed the drastic changes from machine learning to Explainable AI, what could be the future and what is explainable AI in the first place how could it be used and also showing the possible next step such of the explainable AI such as linking probabilistic learning methods with large knowledge representations and logical approaches [14]. In 2023 Sajid Ali and group wrote an article on XAI, which explains the background of XAI and common definitions and how beginners can start with XAI and even shows the evaluation metrics and future research directions [15]. Also in 2023, Katarzyna Borys and group wrote a paper on Explainable AI in medical imaging. Their work aims at establishing common ground for cross-disciplinary

understanding and exchange across disciplines between Deep Learning and healthcare professionals [16]. Also referred to the official SHAP documentation [17] for the implementation of SHAP in supervised learning that is movie success prediction. In 2020 December Pantelis Linardatos and group wrote a research study which focuses on machine learning interpretability methods more specifically a literature review and taxonomy of these methods [18]. In 2020 Matthias Dehmer and group conducted a study which shows the uses and need of explainable AI [19]. In 2019 Feiyu Xu and group published a paper explaining latest progress of Explainable AI at that time and also discussed major research areas and state-of-art approaches done on XAI and ends with discussion of challenges and future directions [20].

3. Research Motivation

The lack of interpretability and explainability is a common issue in predictive models used for movie success prediction, particularly those based on complex algorithms like machine learning and neural networks. To enhance the usefulness and trustworthiness of these models, it is important to develop approaches that not only yield accurate predictions but also provide insights into the factors that contribute to those predictions.

Current movie success prediction models are often based on complex algorithms that are difficult to interpret and explain. This can make it difficult to understand why a particular movie is predicted to be successful or unsuccessful. Developing models that are both accurate and interpretable would make them more useful and trustworthy for decision-making.

Then comes the concept of Explainable AI or XAI. It refers to the development of artificial intelligence models which can provide understandable and transparent explanations for the output of the model. As AI algorithms become more and more sophisticated and pervasive in various domains, we need an increasing need to understand the reason behind the AI prediction or recommendation. Traditional AI models, such as deep learning neural networks, are often opaque, making it difficult to understand how they make decisions. This lack of transparency can be a limitation in critical domains, such as healthcare, finance, and autonomous systems, where trust, accountability, and ethical considerations are important. Explainable AI (XAI) aims to make AI systems more transparent and understandable to humans. By providing explanations for their outputs, XAI can help users, stakeholders, and regulators to understand how AI systems work, identify potential biases or errors, and build trust in these systems.

In this study a basic neural network model was built on the movie success prediction and understand the outputs with the concept of Explainable AI on the model. With this the concept of Explainable AI gets more attention and even

people will start to realize the reliability of the concept and apply it to enhance their predictive models.

4. Methodology

4.1 Dataset and Pre-processing

The first thing in the methodology section starts with the dataset, my dataset is one of the most used and most implemented datasets that is the IMDb dataset. It was obtained from the Kaggle. Kaggle is one of the best platforms for data scientists, machine learning engineers, and researchers to collaborate, share, and compete on data science and machine learning projects. It was founded in 2010 and acquired by Google in 2017. Kaggle provides a diverse range of datasets and challenges for users to explore, analyze, and develop models for various applications.

A lot kind of datasets could have been used such as twitter analysis after some web scraping by releasing the cast and crew information, or director information but it will be biased towards the popularity of the actors like if they are famous it would lead to a lot of positive emotions and the model would think that the popularity would be everything. But it's not the case here, there are some documentaries which have won Oscars, without any actor or a famous director. But what if we go with the plot summaries? For that case, the plot being good or bad would also have a subjective and objective point of views. So even that would be a complex decision. But with this past information which the IMDb dataset has, would provide a lot of information to the model and a lot of patterns can be analysed.

The IMDb dataset originally had 5043 observations to 28 different attributes. It was thought as a reasonable amount of data but after going through the data pre-processing has been done and removed NaN values which were significantly less and removed some unnecessary columns such as movie title and movie's IMDb link.

Then checked correlation of the data as seen in the heatmap in Figure 1, the dataset had two columns causing multicollinearity. In this dataset the attributes causing multicollinearity are `cast_total_facebook_likes` and `num_critic_for_reviews`. Multicollinearity should be eliminated since it decreases the accuracy of computed coefficients, which affects the classification model's statistical strength. Hence removing them would strengthen the model.

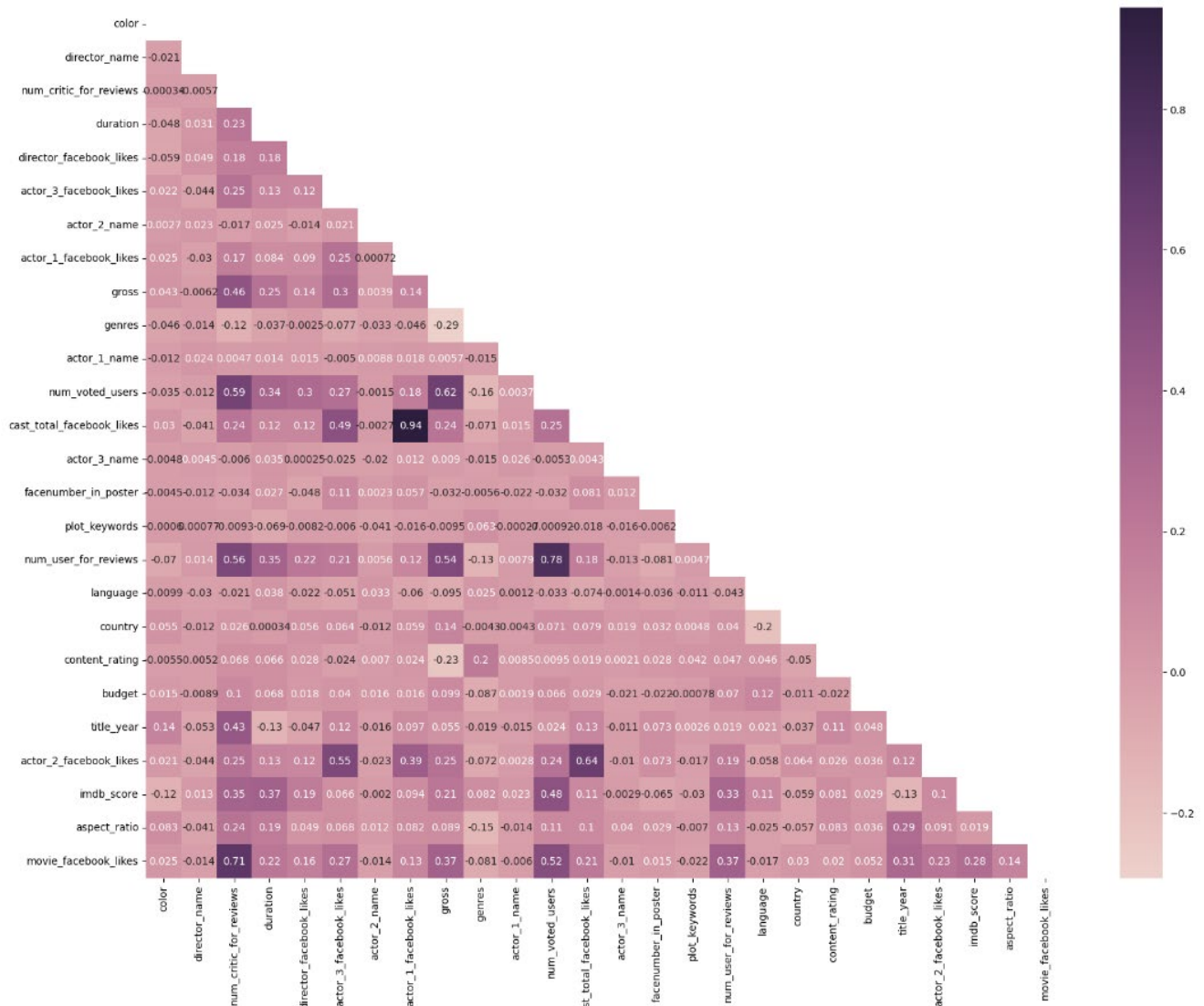


Figure 1. Correlation Heatmap of data

But a classification model would be a very strict than compared to the regression model as it is easy to select a

4.2 Movie Data Labelling

The IMDb dataset is one of the widely used dataset and their website provides a wide range of information about movies and TV shows which also includes cast and crew information, production details and other trivia. The main finding in this website would be the IMDb score. Since IMDb is a popular resource for movie enthusiasts it is one of the go-to source for information about movies and TV-shows and it also provides the users to drop their opinions and reviews on the show along with that the users can also rate the movie/TV-show. Since it is a rating based on a significant population and hence been accepted universally it is good to consider it as one of the best ways to determine the success of a movie. But the IMDb score is in floating point values, and hence it would make the model a regression model.

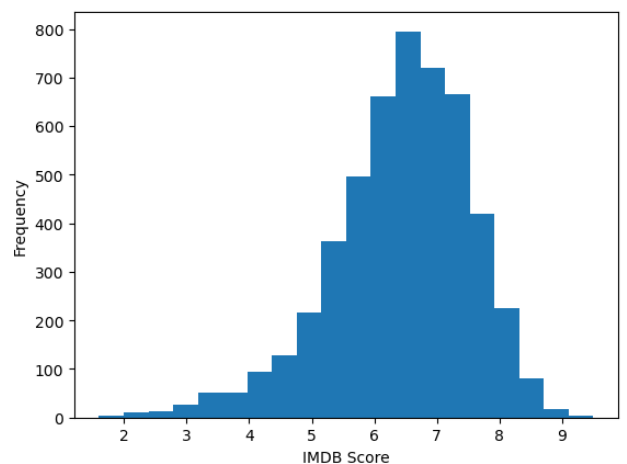


Figure 2. Distribution of IMDb Score

single point value as an output and hence easy to estimate

errors. In Figure 2 you can see that most of the scores are situated around 5.5 to 7.5. After seeing the plot, it was decided to bin the dataset into 3 classes based on their IMDb score namely 'HIT', 'AVG', 'FLOP'. You can see that labelling in Figure 3.

actor_2_facebook_likes	imdb_score	aspect_ratio	movie_facebook_likes	imdb_binned
936.0	7.9	1.78	33000	HIT
5000.0	7.1	2.35	0	AVG
393.0	6.8	2.35	85000	AVG
23000.0	8.5	2.35	164000	HIT

Figure 3. Dataset post labelling

4.3 Proposed model

The suggested framework and a visual representation of the model proposed are discussed in this section. The two main strategies that have been employed to date to forecast movie success were ensemble models and emotive analysis of trailer reviews, both of which had a respectable level of accuracy. And other people used hybrid of deep learning and machine learning or deep learning and ensemble learning. And few used some optimization methods on their hybrid models.

But one of the problems with all of these models we noticed is the complex the model goes the harder it is to reason how the model has come to the specific conclusion for a specific input. Inside the model is a total black box. Consequently, it is known as "black box modelling." Black box models place us in a position where we require a logical justification for whatever output we get from the model. We can employ the idea of explainable AI, which is regarded as a glass-box modelling because it can determine the logical justification for the output or prediction it makes, to resolve this problem. This study a Artificial Neural Network model (ANN) and the SHAP explanations were used. With ANN model we can have the model learn complex patterns in the dataset and combining the model with the features obtained from the SHAP feature importance we can also get the explanations even more clearly. That's why this model was decided to be used.

Explainable AI

The term "explainable AI," or "XAI," describes the capacity of AI models and systems to offer concise and understandable justifications for their choices and behaviours. XAI is particularly important in applications where the consequences of a wrong decision can be significant, such as in medical diagnosis or financial fraud detection. Explainable AI helps us to find out which part of the input is most important for a prediction or also ensures if our model is robust.

SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanation) are the two

most often utilised models. The LIME method is a strategy for emphasising machine learning model predictions. It aims to provide local explanations, meaning explanations that are specific to a particular instance or input. LIME does this by fitting a simpler, interpretable model to the neighbourhood of the instance of interest, and using the coefficients of the simpler model to explain the prediction of the original model.

The black-box model's accompanying predictions are also determined by LIME, which creates a new changed dataset made up of permutated samples. The new adjusted data is then used to train an interpretable model. Models that can be interpreted include decision trees, random forests, decision trees, and logistic regression.

The local importance using LIME method can be evaluated using the below equation (1).

$$f(z) = g(z') + \sum(y - y') * \nabla g(z'). \quad (1)$$

Here,

- $f(z)$ is the prediction of the original model on instance z
- $g(z')$ is the prediction of the simpler model on instance z , which is a perturbed version of z .
- y is a vector containing the features for the original instance z .
- y' is a vector containing the features for the perturbed instance z' .
- $\nabla g(z')$ is the gradient of the simpler model at instance z' .

The equation can be interpreted as follows: the prediction of the original model on instance z is approximated by the prediction of the simpler model on a nearby instance z' , plus a linear combination of the differences between the feature values of z and z' , weighted by the gradients of the simpler model at z' .

The Shapley value is a notion from cooperative game theory that evaluates how much each player in a game contributes to the overall result. The players and the pay-out in machine learning are the input features and the model's output prediction, respectively. In the SHAP technique the Shapley formula plays a huge role in making explanations. The Shapley value formula is employed to calculate the contribution of each feature to the model's final prediction in the equation (2)

$$Shapely_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [modelPred_{S \cup \{i\}}(x_{S \cup \{i\}}) - modelPred_S(x_S)] \quad (2)$$

Here,

- $Shapely_i$ is the Shapley value of feature i
- S is a subset of the set of all features excluding feature i
- $|S|$ is the number of features in subset S
- $modelPred_S$ is the model's prediction using the features from subset S of the input data $modelPred_{S \cup \{i\}}$

represents the model’s prediction based on the input data and features from subset S plus feature i

Here, SHAP and LIME can be used with any machine learning model, regardless of whether it is a neural network, decision tree, or support vector machine because they are both model agnostic in nature. Similarly, methods like cross-validation and grid search are also model agnostic, because they can be used to optimize and evaluate the performance of any machine learning model. Usually SHAP is used in model interpretation, feature selection in for any model generated and get explained values and use them to know what features to use.

SHAP

In this the SHAP concept as the explainable AI addition for my model. In the above section you have seen what is the formula used for the calculation of SHAP values. This section will go somewhat deeper with the concept.

As you know the main idea behind the SHAP or Shapley Additive exPlanations is the cooperative game theory. Imagine a group of people participating in a game, this game has to be a coalition game where people/players form a team to participate in the game, the game can be played cooperatively or non-cooperatively. The best example one could think of is football. And for every goal the team scores they will get a payout from their club. But the second thing here is how would the money they receive is distributed among the team so that there is a fair distribution. And each member of the team has contributed to the goal differently and splitting in some parts will be unfair for some. The answer to this question is called Shapley values, introduced in 1951 by Lloyd Shapley. Shapley values tell us the average contribution of a player to the payout, they fulfill a couple of nice properties and they are the solution yielding a fair distribution. The explainable AI concept SHAP uses this shapley uses this value. Instead of players in a game, relate to features in a machine learning algorithm. Each of this feature contributed differently to a prediction so the prediction would be the payout here and the game would be the machine learning model. That summarizes the basic idea we have treating each feature like a player in a game and calculate shapely values to find out their contribution to the black box model.

The main intuition behind the shapely values is that we want to compare how the correlation would perform with versus without a specific player this way we can find out how this person contributed in the game. Let’s say that a forward in the game scored the goal we’ll give him maximum amount of the payout and the one who passed to him gets the second-place amount. But its not simple as that, we also need to consider the players who distracted defenders and also the players who secured the ball and protected it without letting the other team steal it without their contribution the ball wouldn’t have gone to the forward at all. Summarizing that what we can conclude is we need to consider the interactions between players. So, we need to consider subset of features rather than

considering only single features. And also consider their contributions as well. Shapley values tells us how the prediction is fairly distributed among the individual inputs.

Looking at the formula in the equation (2), the input of this formula looks like $\phi_i(f, x)$ where f is the black box model we are using and x is the input data point and ϕ_i is the shapley value of the feature i . This data point would be a single row in a tabular data set. And tabular data set is must. The data point can be seen as the below Table 1,

$$x = \begin{matrix} A1 = & A2 = & A3 = & A4 = & \dots \\ 456 & 458 & 7898 & 778 & \dots \end{matrix}$$

Table 1. A Tuple from a random dataset with attributes A1, A2, A3 and so on.

In the Table 1, the A1 and A2 and so on are the attributes and their respective values in a tuple x . Now the first thing to do is iterate over all possible subsets S of $F \setminus \{i\}$ so combination of features to make sure to account for the interactions between our individual feature values. For example, the SHAP considers A1, A3. This means only to consider to have information for those two and the values of A2 and other attributes/features are unknown. And now the most important step we get the black box model output for this subset with and without the feature we are interested in. The difference in those two tells us how the attribute we are interested in contributed in the prediction of the subset. The difference mention earlier is here $modelPred_{S \cup \{i\}}(x_{S \cup \{i\}}) - modelPred_S(x_S)$ where $modelPred_{S \cup \{i\}}(x_{S \cup \{i\}})$ represents the prediction with the attribute we are interested in $\{i\}$ and $modelPred_S(x_S)$ represents the prediction without the attribute. That is also called the marginal value and this step we do it for each possible combination of attributes and subsets and each of those are additionally weighted according to how many players are there in the correlation, in other terms how many features and the total number of features are there in the subset.

$$\frac{|S|!(|F|-|S|-1)!}{|F|!} \text{ so, } |F| \text{ is the total number of features, the}$$

intuition is that contribution of adding the feature A1 should be weighted more if already many features are included in the subset. So that would tell us that this specific feature gives us a strong change in the prediction even if many other features are already included. On the other hand, we also want to give more weight to the small correlations, because there we have the features isolated and we can directly observe their effect directly on the predictions.

However, there is one more question how do we exclude a feature from a machine learning model, typically the inputs are fixed size and we cannot just remove parts of it because then the shape would change. The way this is solved in SHAP is that for the features we want to exclude we just input random values from the training dataset, and random feature value has not much worth to contribute in the prediction.

Artificial Neural Network

In this paper, an artificial neural-network (ANN) is utilized to intake the features from the tabular data and then predict if the movie will become successful or not. A neural network is a machine learning technique based on the human brain which is comprised of a complex system of

In the above equation (3):

- HLA_k^{i+1} is the output value after activating of the $(i + 1)^{th}$ layer of k^{th} neuron
- $Weight_{jk}^{i,i+1}$ is the weight assigned to the edge between

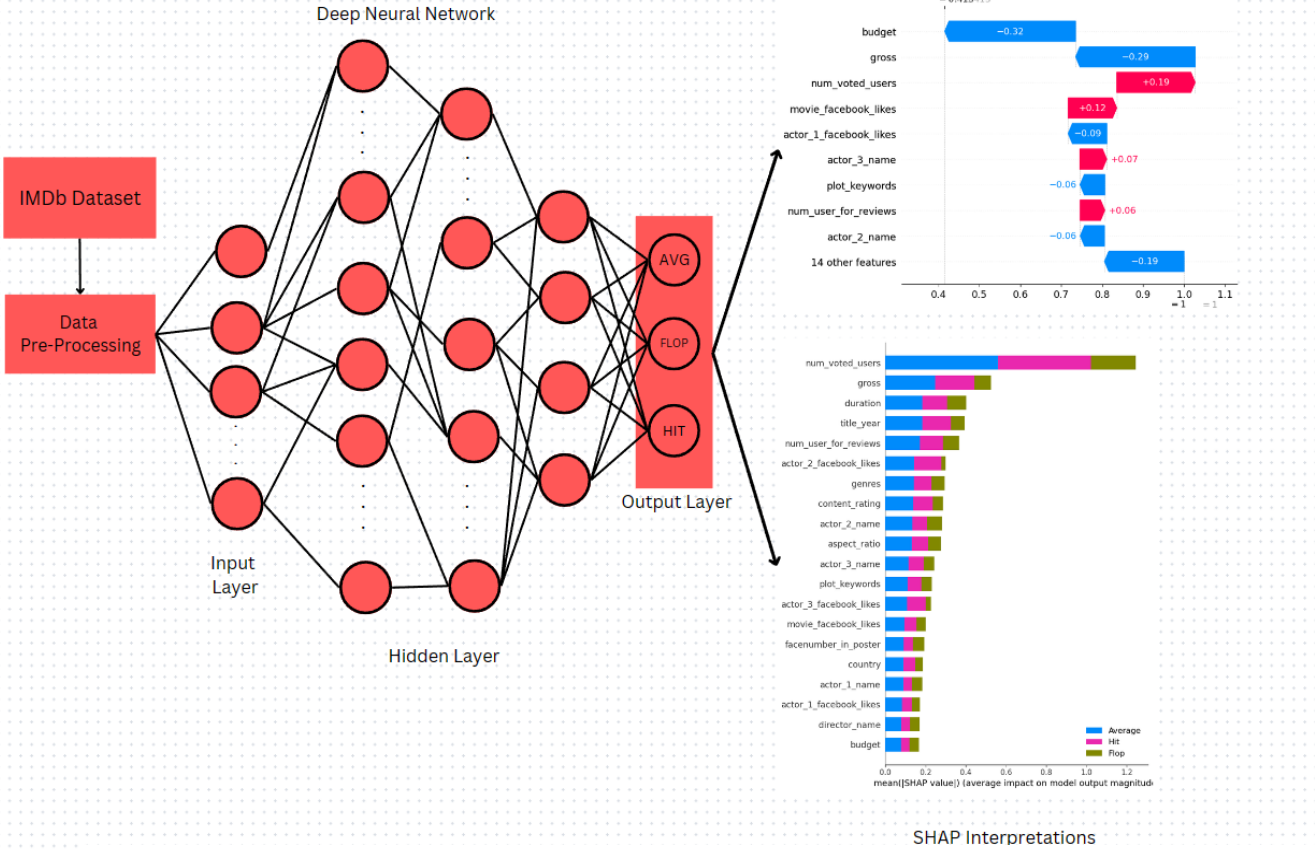


Figure 4. Model Architecture

interconnected neurons, which work together to process information, in a way that allows the network to learn from data and make predictions or classifications. An input layer, a hidden layer, and an output layer are the key elements of an ANN.

Input Layer of ANN

The input layer is the first layer of the neural network, that receive the input data from the dataset. From here the input data is transmitted into neurons and then passed on to the next layer.

Hidden Layer of ANN.

The hidden layer, which is made up of one or more layers of neurons, sits between the input layer and the output layer. In hidden layer artificial neurons take in a set of weighted inputs and produce an output through an activation function. We can have one or more hidden layers in a neural network model.

A neuron in hidden layer can be activated as represented below:

$$HLA_k^{i+1} = Activation_Function(\sum_j HLA_j^i * Weight_{jk}^{i,i+1} + bias_k^{i+1}) \quad (3)$$

the j^{th} neuron in the i^{th} layer and k^{th} neuron in the $(i + 1)^{th}$ layer.

- $bias_k^{i+1}$ is the bias of the k^{th} neuron present in the $(i + 1)^{th}$ layer

In this paper, the activation function was the ReLu function. It is indicated that the ReLu activation function as in the equation (4):

$$ReLU(y) = Max(0, y) \quad (4)$$

Here $Max(0, y)$ stands for Maximum of $(0, y)$.

Output Layer

The output layer is the final layer of the neural network, and it is responsible for the final predictions made by the model. The number of neurons in the output layer is determined by the number of classes or the type of problem being solved. For this article since it was decided to split the dataset into 3 classes the output variable will have 3 neurons.

The output layer takes in the inputs from the last hidden layer and performs calculations similar to the hidden layer. But here the activation function used is different compared

to the hidden layers. Here SoftMax activation function was used to compute probabilities of each class and those probabilities determine the prediction of the model. The SoftMax function is represented as in equation (5):

$$SoftMax(Y = y_i|X) = \frac{e^{(h_i)}}{\sum_k e^{(h_k)}} \quad (5)$$

The visual representation of the model architecture is shown by the Figure 4

5. Results and Discussion

This section mainly deals with the performance evaluation of the model. This model's performance metrics are recall, accuracy, f1-score, and precision. The above four methods are few best methods to evaluate a classification model. The percentage of cases when HIT, AVG, and FLOP are

model training. Accuracy obtained is 75%, F1-score is 83% for Predicting Averages and 41 for predicting Flops and 60% for predicting Hits. Recall is 87% for predicting Averages, 33% for predicting Flops and 57% for predicting Hits. And Precision is 79% for predicting averages 54% for predicting flops and 64% for predicting Hits.

After applying SHAP we can try to interpret the local explanation for some random sample it can seen in the results in Figure 5,6,7. The result of the prediction was AVG and hence the SoftMax value was given 0.99 for the AVG. And it can be seen that how much an attribute contributes in shifting the value such as budget and gross estimate in the above local explanation.

A summary of the SHAP values produced using the suggested model can be represented by Figure 8. From the plot it can be interpreted that our top features are num_voted_users, gross, title_year, content_rating. The global explanation can be interpreted from Figure 8. The

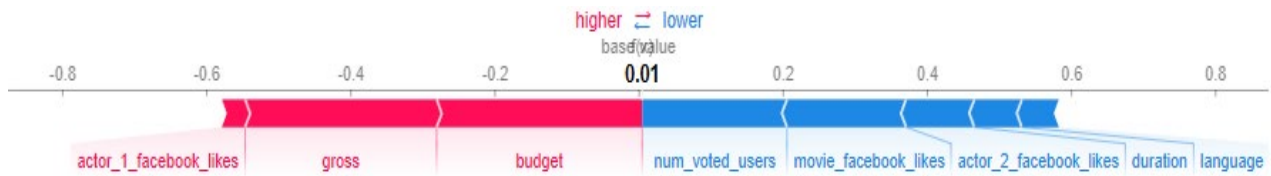


Figure 5. Local interpretation for the HIT value



Figure 6. Local Interpretation for the FLOP value

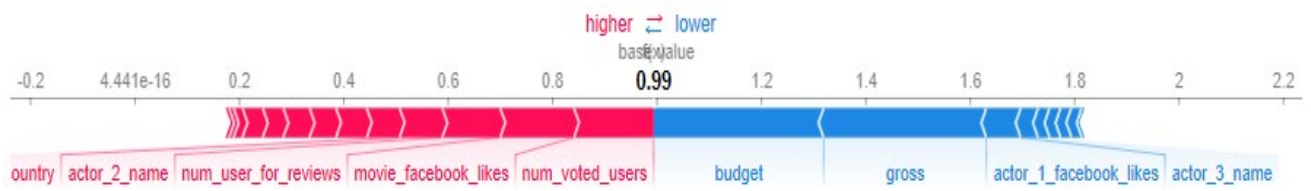


Figure 7. Local Interpretation for the AVG value

properly identified to the entire test set serves as a measure of accuracy. The ratio of cases that were classified as HIT, AVG, and FLOP to all instances that were classified in that manner is known as precision. Recall can be described as the proportion of cases that were classified as HIT, AVG, or FLOP to all instances that belonged to that category. The harmonic mean of recall and precision is the F1 score.

An ANN model implemented with Keras library using Sequential method having learning rate of 0.001, 60 epochs, and a batch size of 32 was used for the suggested

colors represent the attribute, and the length of the bar represent how much the attribute influences on the overall dataset. The longer the bar the more influence and vice versa. It can be noticed that num_voted user has the most influence on all the classes. This explains that the greater number of voters can easily influence the IMDb score and hence it means the movie is more watched and is well liked by the watchers

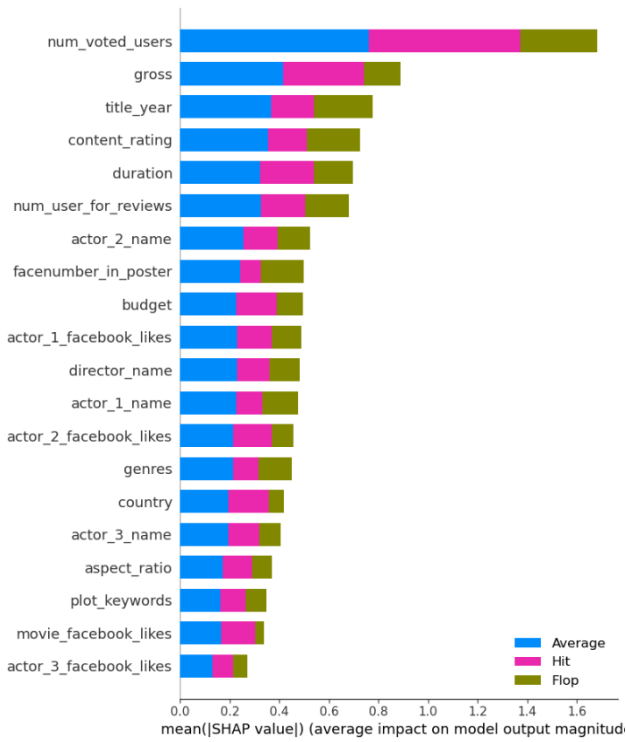


Figure 8. Distribution of SHAP values over the dataset.

	precision	recall	f1-score	support
0	0.80	0.86	0.83	788
1	0.43	0.30	0.35	149
2	0.66	0.60	0.63	190
accuracy			0.74	1127
macro avg	0.63	0.59	0.60	1127
weighted avg	0.72	0.74	0.73	1127

Figure 9. Classification Report of ANN model

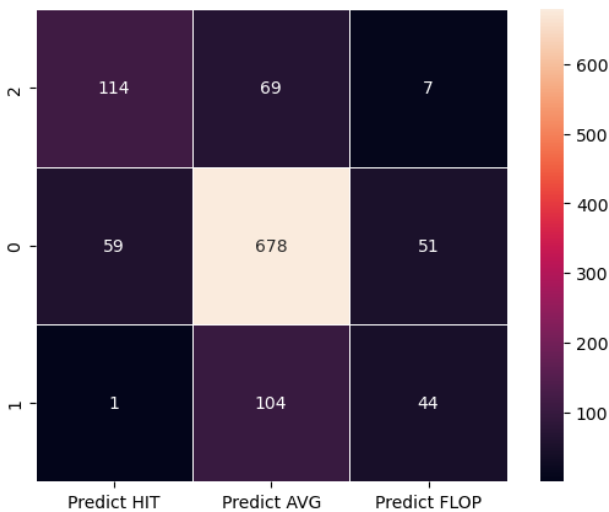


Figure 10. Confusion Matrix of ANN model

Figure 9 and Figure 10 shows us the classification report and confusion matrix, there's a lot of gap in the number of AVG samples and FLOP samples and HIT samples so the model is sometimes predicting some HIT movies and as well as FLOP movies as AVG. It is due to lack of enough movies in HIT and FLOP range.

6. Conclusion and Future Work

In this study, an ANN model is used to analyse and study the factors which influence the success of a movie. Neural network uses complex calculations and feature extractions to gain a higher accuracy. But when researchers analyse and make various predictions sometimes due to model being complex they can't find an explanation for all the predictions made by the model especially when it comes to Neural Network models. Understanding the reasoning behind a certain prediction is crucial for improving the model's ability to be explained. That's why SHAP values were used to get the local interpretations and global interpretations.; these will be helpful to understand what features helped our model to make that prediction. But when it comes to the accuracy of the model we built it is not up to the mark to what other researchers have done but it is somewhat improvised compared to the existing machine learning techniques. But finally, it can be explained what features give importance to FLOP region or HIT region.

But In future aspects, additional methods will be implied on the same data or various data will be collected and new models will be applied and predictions will be made so as to make more accurate predictions and explanations.

7. References

- [1] W. Zhang and S. Skiena. Improving movie gross prediction through news analysis. In *Web Intelligence*, pages 301-304, 2009.
- [2] Michael T. Lash and Kang Zhao, Early Predictions of Movie Success: the Who, What, and When of Profitability, June 2015.
- [3] You Jin Kim, Yun Gyung Cheong, and Jung Hoon Lee. 2019. Prediction of a Movie's Success From Plot Summaries Using Deep Learning Models. In *Proceedings of the Second Workshop on Storytelling*, pages 127–135, Florence, Italy. Association for Computational Linguistics.
- [4] Sahu, S., Kumar, R., Long, H.V. et al. Early-production stage prediction of movies success using K-fold hybrid deep ensemble learning model. *Multimed Tools Appl* 82, 4031–4061 (2023).
- [5] Ni, Yuan, et al. "Movie Box Office Prediction Based on Multi-Model Ensembles." *Information* 13.6 (2022): 299.
- [6] Wei Lu, Xiaoqiao Zhang, Xinchen Zhan, "Movie Box Office Prediction Based on IFOA-GRNN", *Discrete Dynamics in Nature and Society*, vol. 2022, Article ID 3690077, 10 pages, 2022.
- [7] G. Velingkar, R. Varadarajan, S. Lanka and A. K. M, "Movie Box-Office Success Prediction Using Machine Learning," 2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T), Raipur, India, 2022, pp. 1-6, doi: 10.1109/ICPC2T53885.2022.9776798.
- [8] Pande, S., & Khamparia, A. (2023). EXPLAINABLE DEEP NEURAL NETWORK BASED ANALYSIS ON INTRUSION DETECTION SYSTEMS. *Computer Science*, 24(1).
- [9] Guo, M.; Yuan, Z.; Janson, B.; Peng, Y.; Yang, Y.; Wang, W. Older Pedestrian Traffic Crashes Severity Analysis Based on an Emerging Machine Learning XGBoost. *Sustainability* 2021, 13, 926.
- [10] Mokhtari, Karim El, Ben Peachey Higdon, and Ayşe Başar. "Interpreting financial time series with SHAP values." *Proceedings of the 29th annual international conference on computer science and software engineering*. 2019.
- [11] [arXiv:1903.02407v2](https://arxiv.org/abs/1903.02407v2) [cs.LG]
- [12] Främling, K. (2020). Decision Theory Meets Explainable AI. *Explainable, Transparent Autonomous Agents and Multi-Agent Systems*, 12175, 57 - 74. https://doi.org/10.1007/978-3-030-51924-7_4.
- [13] Fernandez A., Herrera F., Cordon O., del Jesus M.J., Marcelloni F.: Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to?, *IEEE Computational Intelligence Magazine*, vol. 14(1), pp. 69–81, 2019.
- [14] Holzinger A.: From machine learning to explainable AI. In: 2018 world symposium on digital intelligence for systems and machines (DISA), pp. 55–66, IEEE, 2018.
- [15] Ali, Sajid, et al. "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence." *Information Fusion* (2023): 101805.
- [16] Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners—Beyond saliency-based XAI approaches." *European journal of radiology* (2023): 110786.
- [17] [Welcome to the SHAP documentation — SHAP latest documentation](#)
- [18] Linardatos, Pantelis, Vasilis Papastefanopoulos, and Sotiris Kotsiantis. "Explainable ai: A review of machine learning interpretability methods." *Entropy* 23.1 (2020): 18.
- [19] Emmert-Streib, Frank, Olli Yli-Harja, and Matthias Dehmer. "Explainable artificial intelligence and machine learning: A reality rooted perspective." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10.6 (2020): e1368.
- [20] Xu, Feiyu, et al. "Explainable AI: A brief survey on history, research areas, approaches and challenges." *Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II* 8. Springer International Publishing, 2019.