ANALYSIS AUDIENCE ACCEPTANCE BASED ON TYPE USING LOGISTIC REGRESSION

Dr.M.Rajeswari,
Assistant Professor,
Department of B.Com (A&F and BA),
PSGR Krishnammal College for Women, Coimbatore, India.
rajeshwarim@psgrkcw.ac.in

N.Dhanya,
UG Scholar,
Department of B.Com (Business Analytics),
PSGR Krishnammal College for Women, Coimbatore, India.
Dhanyanandhakumar2001@gmail.com

Abstract:

This project is about analysing the audience acceptance which genres or type movie is highly accepted and succeed on the view of audience. By using type and certificate, predicting whether this type of genres succeed or failure on the basis of audience view.

Keywords: Logistic Regression, succeed, type.

I. INTRODUCTION:

By using machine learning algorithms analysing audience acceptance. For that, we used an analysis tool ' JUPYTER NOTEBOOK IN PYTHON'. It contains tools for data pre-processing, transformation, removing null values and regression. Here, logistic regression is to use to analysing success or failure (which genres highly succeed and accepted by the audience). The success of the movie is based on a budget, director, type, title and story of the movie. If Movie is not succeed, maybe make a change in budget or concept of the movie it maybe get success. We will classify the movie as hit, flop, super hit on the basis of investment made by the movie for example if the movie was made on budget 50 crores and it has collected 40 crores on Indian box office then it will be like losing the movie even if it has earned worldwide. The prediction made is categorized into three flop, hit and super hit. Normally the net income defines the terms whether the movie is hit, flop, super hit.

1. Flop: There are a lot of points depending on which flop is declared. Using different parameters gives different results. Many points are calculated for the same such as box office collection, net incomes. If the movie was made on the budget of 50 crores and if the movie has not earned maximum revenue on Indian box office, then it will predict losing the movie.

2. Hit: If the movie earns the profit 20 percent more than the budget than the movie predicted hit.

3. Super Hit: For the movie to be a super hit budget plays a vital role. If the movie earns 50% more than the budget than the movie predicted as the super hit [2].

II. OBJECTIVE

Analyse success and failure of movie type and also analyse audience acceptance which movie type highly like by audience.

- Find Success and failure of movie type
- Prefer the solution
- Find which movie highly view by audience.

III. RELASED WORK

Now days, Movie is good source of entertainment and stress relief show. The each and every movie is depend on the rating of people. The rating of the movie are two methods the first is 1 to 10 rating method and another one star rating methods. By analysing database, the first thing that we would like to know is how many movies there are in each category. Then only we predict the movie is highly liked by the viewers. Movie recommendation systems usually predict what movies a user will like based on the attributes present in previously liked movies [9].

Making a prediction of any movie in sense of popularity is an emerging field of analysis. Film industry has been growing rapidly in last 20 years in all over world. A huge volume of data related to movie is available on web. Different kinds of movies released every year. Some movies inspired and motivate us to do something new in our life, some movies make us happy, some of them are for fun and entertainment, some movies portray historical event while some do may more [7].

A good amount of study has been carried out by various authors for movie popularity prediction based on data available on IMDb. Success of any movie primarily depends on numerous factors like story, genre, director, actors/actresses etc.([4], [1],

[6]. These models prediction were done by usind IMDb movie data source. set up a predicted the movie success based on the revenue generated by movie. The work was explored using volume of Twitter sentiments on sample of 24 movies. They achieved an adjusted coefficient of determination of 0.97 before the movie release for the first weekend revenue [8].

However, Logistics regression (LR) is one of the most popular algorithm that can be used to predict a binary outcome and one or multiple continuous or categorical predictor variables. In fact, LR can be used for outcome variable with more two categories as well [2]. Since LR uses logit model therefore the value of outcome variable ranges only between 0 and 1. More importantly, LR does not follow the assumption of normality of sample data thus making it much more flexible to use it for prediction. [5].

IV. METHOLOGY

LOGISTIC REGRESSION

Logistic regression is that the appropriate multivariate analysis to conduct when the variable is dichotomous (binary). Like all regression analyses, the logistic regression may be a predictive analysis. Logistic regression is used to elucidate data and to elucidate the connection between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.

FLOW CHART
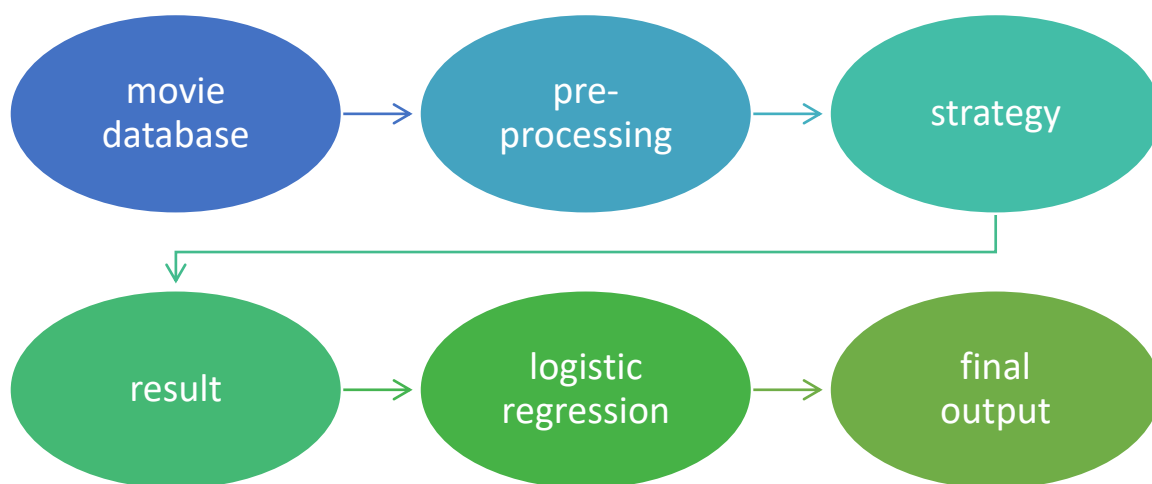
V. RESULT:

Pwd to check the directory. Then import the libraries package like numpy, pandas, matplotlib, logistic regression for analysing the database, pyplot for importing numbers and displaying the graph. To import the database and fix the current location **C:\Users\dhany\jupyter.** In logistic regression to split training and test database into 80% and 20%. in logistic regression the train_test_split is import from sklearn. **Data.head()** is used to show the data in database and head() is used to first n rows. **Data.columns()** is used to used to return the column values. **Data.tail()** is return last n rows. **Data.describe()** is used to show some basic statistical details like mean, count, standard deviation, minimum, maximum etc. **print(data.isnull().sum())** it shows how many null values in database. **print(data.shape)** is used show number of row and columns count.
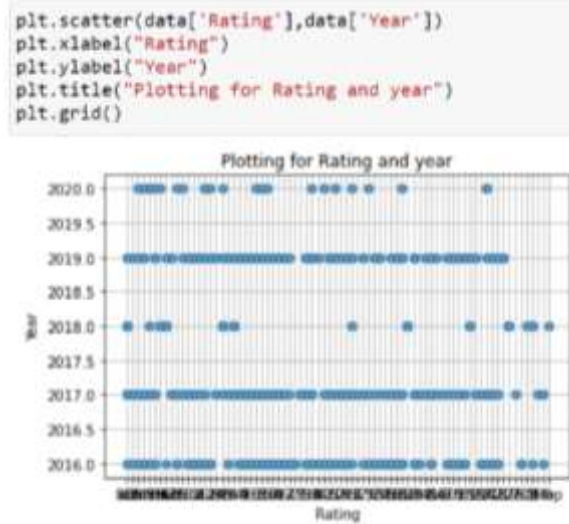
```
plt.scatter(data['Rating'],data['Year'])
plt.xlabel("Rating")
plt.ylabel("Year")
plt.title("Plotting for Rating and year")
plt.grid()
```


Fig: 1.1

```
plt.scatter(data['Certificate'],data['Type'])
plt.xlabel("Certificate")
plt.ylabel("Type")
plt.title("Plotting for type and certificate ")
plt.grid()
```


Fig:1.2

In above fig:1.1,fig:1.2, By analysing the scatter plot , most of the audience highly prefer commerical movie type and U/A and U Certificate.

```
prediction_var=['No_of_Votes','Year']
```

```
train_X=train[prediction_var]
train_Y=train['Title']
```

```
train_X=train[prediction_var]
train_Y1=train['Type']
```

```
train_X=train[prediction_var]
train_Z=train['Directors']
```

```
test_X=test[prediction_var]
test_Y=test['Title']
```

```
test_X=test[prediction_var]
test_Y1=test['Type']
```

```
test_X=test[prediction_var]
test_Z=test['Directors']
```

fig:1.3

Fig:1.4

In fig:1.3fig:1.4 Train and test variable Genres, title, directors. Movie maybe succeed and accepted by audiences on basic Genres, title, Directors. Each and every person prepare different type(love,Horror,thriller,commerial)and different Genre(Drama,comedy,crime,romance,horror). In fig:8.4 using logistic regression to analysis the audience acceptence using type, Genre and rating of the movie.By using the strategy visualisation, analysing database for the purpose of which type or Genre movie is highly view by audience. In strategy one, we used tilte,type,Verdict and strategy two we used Title,Director,Verdict. In Strategy three we used Title,cast,Verdict. In Strategy four we used Title,Budget,Verdict and strategy five we used Title,Releaseddate,Verdict. They used blockbuster strategy to win the verdict flop movies.

If movie earns the profit 25 percent more than the budget predicted movie hit. If movie earns the profit 50 percent more than the budget predicted movie as superhit.if movie earns the less than the budget movie as flop. f the movie hit or flop or superhit it is based on the genres, type and director and title of the movie. By using trailer rating they can identify whether the movie get flop or hit, superhit . They identify movie going to get flop means they can make a changes in budget.

VI. CONCLUSION AND FURTHER WORK

In this paper,python used to analyse the success of movie and which movie is highly accepted by the audience. The main intention of the paper, perparing the solution for failure movie. We used 700 data which contains the various factors like box office, rating, tailor rating, Genre, directors etc.

VII. REFERENCE

[1] A. Chen, "Forecasting gross revenues at the movie box office," Working paper, University of Washington, Seattle, WA, June 2002.

[2] D. G. Kleinbaum, and M. Klein, Logistic Regression, New York, Springer, 2010, pp. 1-39.

[3] International Journal Of Engineering Research And Development e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.comVolume 16, Issue 6 (June 2020), PP. 10-14.

[4] J. S. Simonoff and I. R. Sparrow, "Predicting Movie Grosses: Winners and Losers, Blockbusters and Sleepers," Chance, vol. 13, no. 3, pp. 15–24, 2000.

[5] M. Pohar, M. Blas, and S. Turk, Comparison of logistic regression and linear discriminant analysis, Metodoloki zvezki, Vol. 1(1), 2004, pp. 143-161

[6] M. S. Sawhney and J. Eliashberg, "A Parsimonious Model for forecasting Gross Box-Office Revenues of Motion Pictures," Marketing Science, vol. 15, no. 2, pp. 113–131, 1996.

[7] Prediction and Analysis of Movie Success with Machine Learning Approach February 2020

[8] S. Asur, BA. Huberman," Predicting the future with social media. In: Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology", pp. 492-499, 2010

[9] srs reddy, sravani nalluri, subramanyam kunisetti, s. ashokand, b. venkatesh :Content-Based Movie recommendation system Using Genre Correlation.