**Movie Success Prediction Using Naïve Bayes, Logistic Regression and Support Vector Machine**

**Abstract** :

The entertainment industry is a rapidly growing billion-dollar industry. With new milestones being reached almost every day, this industry has proved itself to be a very profitable business, if done correctly. Since huge investments are involved in the production and making of movies, both in terms of time and money, it would only make sense to try to predict the outcome beforehand. In an attempt to tackle this problem, we have built a model that predicts whether or not a movie can be called a success. The model compares the performance of three machine learning algorithms i.e. Naive Bayes, Logistic Regression, and Support Vector Machine (SVM), over two different datasets, to observe which performs better. We have illustrated the model, as well as its results, findings, and observations in this literature.

**EXISTING SYSTEM:**

The prediction of success of movie with good accuracy is needed in the film industry which helps different people working in the film industry manly for the investors it is one of the major sources of global commerce and marketing. Movies create a new craze among people especially young people. Not only movie directors and box office officials are concerned with the success of the movie but general people also. People used to talk about these in social medias. Therefore, analysis of social media data about the movies is popular among the data analysts. And remains some other scopes like analyzing a director’s previous success histories or a actor’s previous popularity etc. Again, the analysis may be different on different countries. Naturally peoples from all the regions of the world do not react in the similar way. Movies are now available on internet. There are platforms like IMDb (Internet Movie Database) , Rotten Tomatoes , Meta critics etc. where people can share their reviews about movies. Movies continue to be a major source of entertainment in any country.

**DISADVANTAGES:**

* Less Accuracy
* Security is less.

**PROPOSED SYSTEM:**

We proposed to develop a model for predicting the success of movie being a Flop or Hit , before a movie is actually released using machine learning techniques and algorithms We have examined three different methods of film classification based on their ability to predict a film's commercial success. Based on experimental results and the collected dataset, it appears that the Logistic regression classifier outperforms both the C4.5 and the Naive Bayes classifiers. More characteristics and a larger dataset will allow for more precise predictions in the future. With the aid of such frameworks, an automated tool can be created, which may be useful for movie suggestion to consumers based upon rating of success in the sense of popularity of any movie. In addition, by incorporating extra elements in the form of hybrid approaches, such a system may be expanded to a recommender system. The opinions of users can also be included in to improve the efficiency and accuracy with which movies are predicted.

**ADVANTAGES:**

* High Accuracy
* Security is high
* LIdar (light detection and ranging), also known as 3D laser scanning, is a tool that self-driving cars use to scan their environments with lasers. A typical lidar sensor pulses thousands of beams of infrared laser light into its surroundings and waits for the beams to reflect off environmental features.

**HARDWARE REQUIREMENTS:**

# Processor - Intel i3 or i5 (min)

* Speed - 1.1 Ghz
* RAM - 4GB(min)
* Hard Disk - 500 GB(min)
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* Operating System - Windows10 (min)
* Programming Language - Python

|  |
| --- |
|  |
|  |
|  |
|  |