iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** XII **Month of publication:** December 2021

DOI: https://doi.org/10.22214/ijraset.2021.39206

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XII Dec 2021- Available at www.ijraset.com

TV Show Popularity Analysis Using Data Mining

Deesha Rawat[1], Gracy Parkhe[2], Devashish Bhand[3]
Affiliation: Acropolis Institute of Technology And Research

Abstract: Evaluating and updating TV programs using data mining. Guessing the popularity of the TV show is an exciting activity with a growing interest in people in TV dramas. A simple guess of a trending TV game based on individual ratings can be made. Simple predictions of trending TV shows based on individual ratings can be made based on the target audience (Age, Year of Release, Rotten Tomatoes, etc.).
Keywords: TV Show, Data Mining, Popularity Analysis, Data Visualization, IMDB Dataset

I. INTRODUCTION

Evaluating and rating TV shows using data mining. Guessing the popularity of the TV show is an exciting activity with a growing interest in people in TV shows. A simple guess of a trending TV show based on individual ratings can be made. Simple predictions of a trending TV show based on individual ratings can be made based on the target audience (Age, Year of Release, Rotten Tomatoes, etc.). This project uses machine learning to compare TV Shows to analyze their popularity. A set of data is categorized and calculated show views, based on limited display views. Once the rating is done their comparisons are made using data visualization, ratings for different television shows can be viewed using bar graphs, pie charts, and histograms.

II. PROBLEM STATEMENT

It is necessary to create a prediction model that acts as a rating system that will be useful to those who are willing to watch a particular show. They will also be able to get feedback from previous audiences and provide support for new audiences. A large website of online websites and social networking sites will be used for the same. TV shows have gained more popularity among people than anything else in the last decade.

III. OBJECTIVE

A. To show a comparison between different TV shows.
B. Visualization of the data to find useful insights.
C. Developing an interactive user interface

IV. LITERATURE REVIEW

TV shows are thriving worldwide each day. The popularity of the TV show using data mining. A simple guess of a trending TV game based on individual ratings can be made. TV shows will be rated based on the number of viewers. The more you look, the more they will earn. Then their comparisons can be viewed in the form of graphs using data visualization.

A. Current System

1) *Motion picture Success Prediction utilizing Data Mining:* The systems were connected to the film website and therefore the data met the improvement and intermixture method before drilling ways were used. the knowledge mine manages samples of the knowledge provided and detects patterns. It detects hidden patterns and connections between totally different objects. Such communication will separate chronology, coming up with an assortment. They provided a very important model for this investigation that would cut back the probabilities of disappointment and supply colleagues with a transparent assurance and prediction of progress.
2) *Group of onlookers Rating Prediction of TV Dramas dependent on the Cast*: Yusuke Fukushima, Toshihiko Yamasaki and Kiyoharu Aizawa of the University of Yedo in 2016 took a firearm at a check to check the performance of the sport before communication while not victimization video or a bunch of viewers' reaction to them. Considering the flow of featured artists enclosed, the daily comments on Wikipedia of screen actors and their posts on Twitter were used as symbols of information. it had been clear that the show gets a high rating once there square measure totally different known characters on screen.
3) *Limitation*: The limitation is that these forums have one feature. Some platforms even need funding to use a number of options. to realize such options, viewers should purchase them.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XII Dec 2021- Available at www.ijraset.com

V. METHODOLOGY

The methodology includes the steps to be followed to achieve the objective of the project during the project development.

1) Gathering data from IMDB data.
2) Cleaning data to have homogeneity.
3) Model building, selecting the right algorithm.
4) Gaining insights from the model results.
5) Data visualization - transforming data into a graph

A. Project Flow Diagram

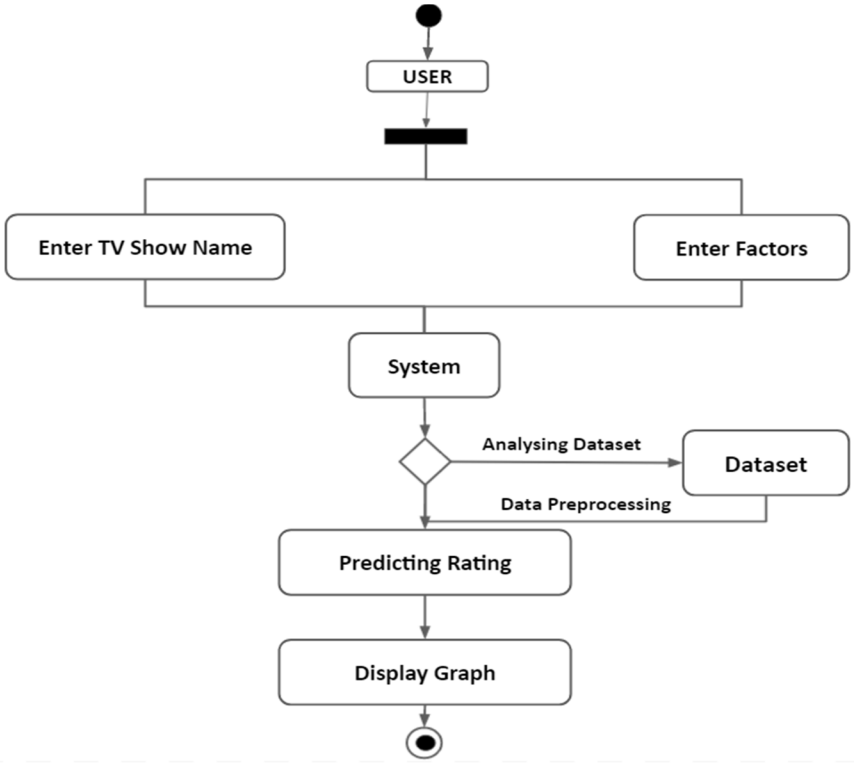


Fig 1: Activity Diagram

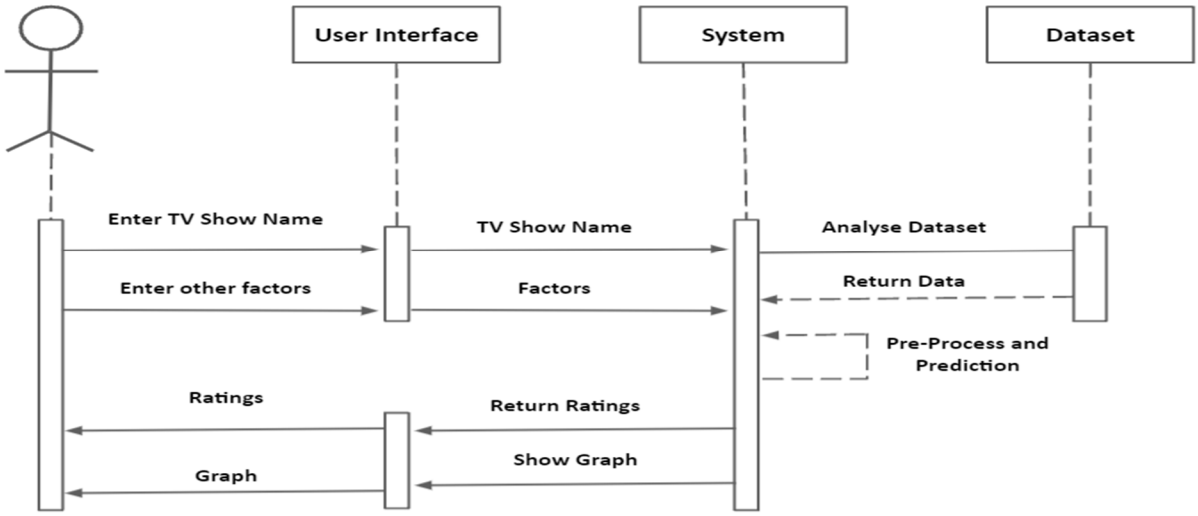


Fig 2: Sequence Diagram

©IJRASET: All Rights are Reserved

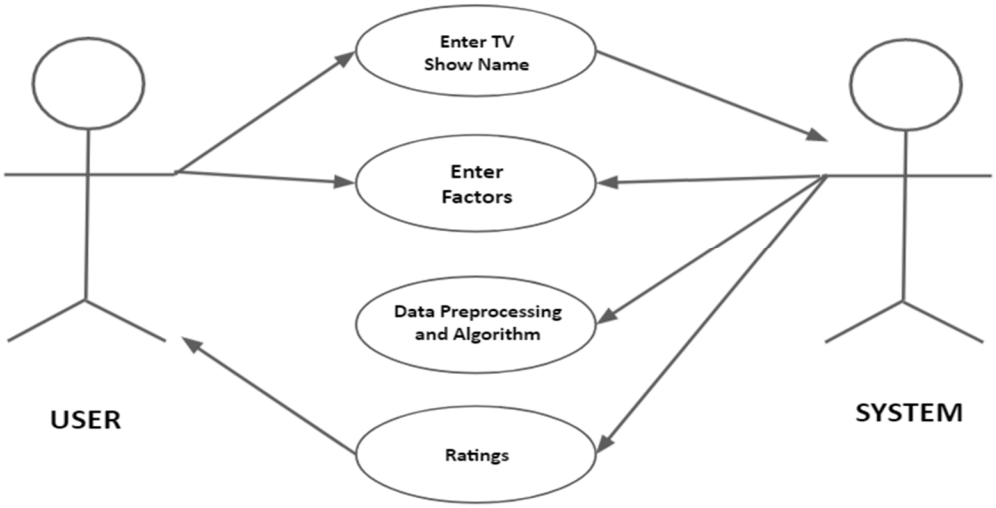


Fig 3: Use Case Diagram

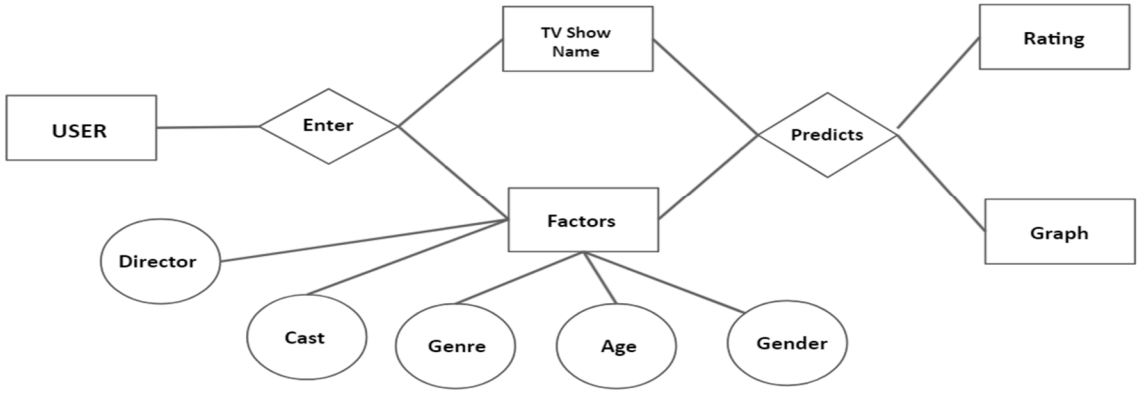


Fig 4: ER Diagram

B. Implementation Images



Fig 1: TV Show Title WorldCloud 100 Words

©IJRASET: All Rights are Reserved

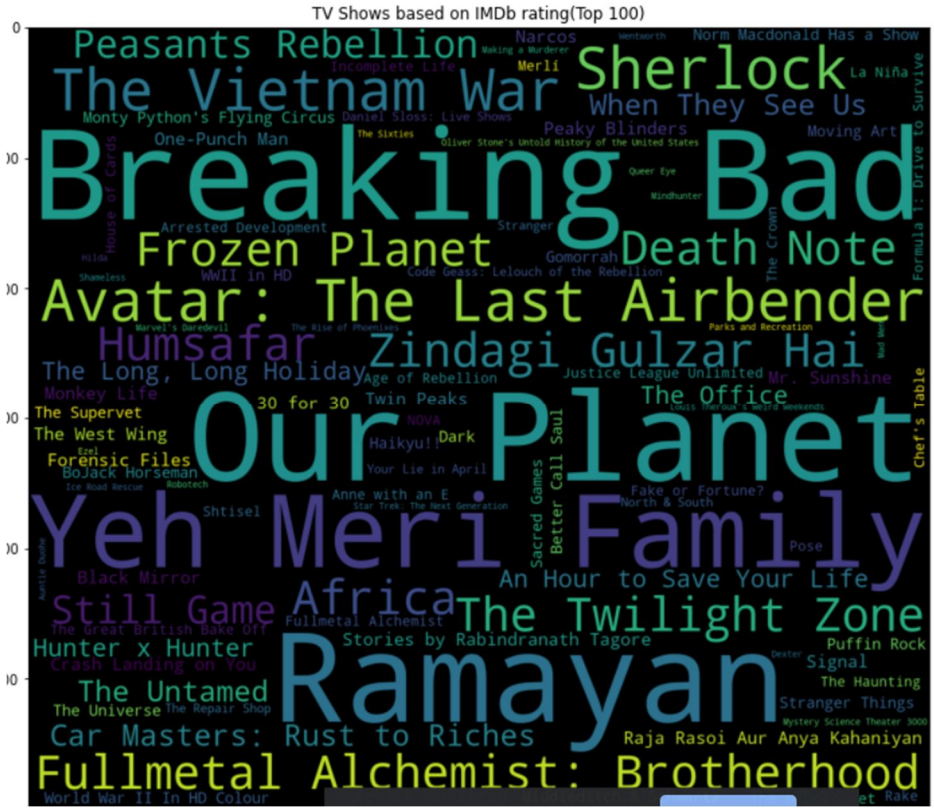


Fig 2: TV Shows based on IMDB rating

```
#overall year of release analysis

plt.subplots(figsize=(8,6))
sns.distplot(data["Age"],kde=False, color="red")
```

```
<AxesSubplot:xlabel='Age'>
```

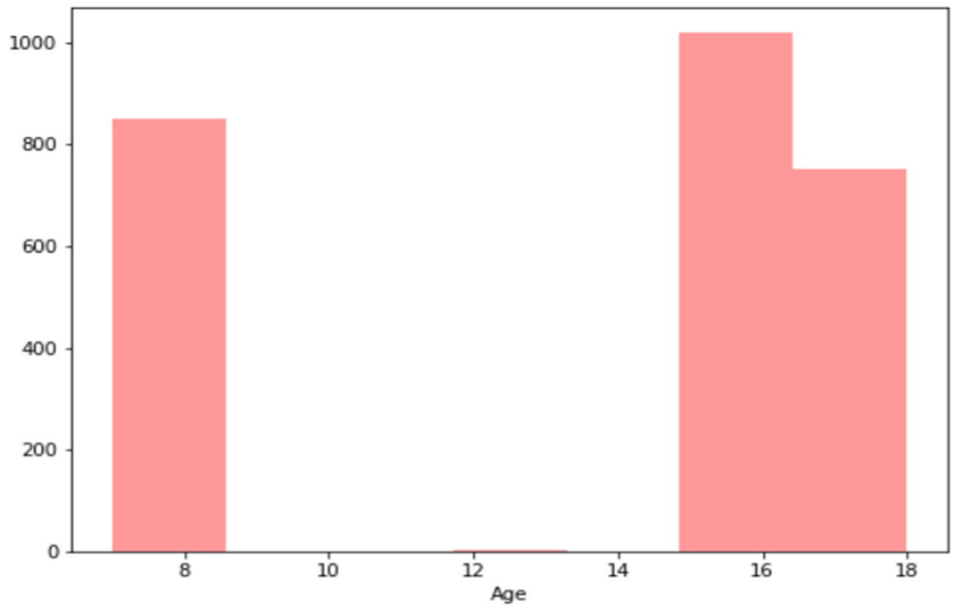


Fig 3:Analysis of Year Releases

©IJRASET: All Rights are Reserved

```
print("TV Shows with highest IMDb ratings are= ")
print((data.sort_values("IMDb",ascending=False).head(20))['Title'])
```

```
TV Shows with highest IMDb ratings are=
3023                              Destiny
0                           Breaking Bad
3747                        Malgudi Days
3177                        Hungry Henry
3567                    Band of Brothers
2365                 The Joy of Painting
4128                      Green Paradise
91                           Our Planet
3566                           The Wire
325                             Ramayan
1931                      Rick and Morty
4041                    Everyday Driver
3701                           Baseball
282                     Yeh Meri Family
3798                            The Bay
4257                 Single and Anxious
3568                       The Sopranos
4029            Harmony with A R Rahman
9             Avatar: The Last Airbender
15      Fullmetal Alchemist: Brotherhood
Name: Title, dtype: object
```

Fig 4: List of TV Shows with highest IMDB ratings

```
#barplot of rating
plt.subplots(figsize=(8,6))
sns.barplot(x="IMDb", y="Title" , data= data.sort_values("IMDb",ascending=False).head(20))
```

```
<AxesSubplot:xlabel='IMDb', ylabel='Title'>
```

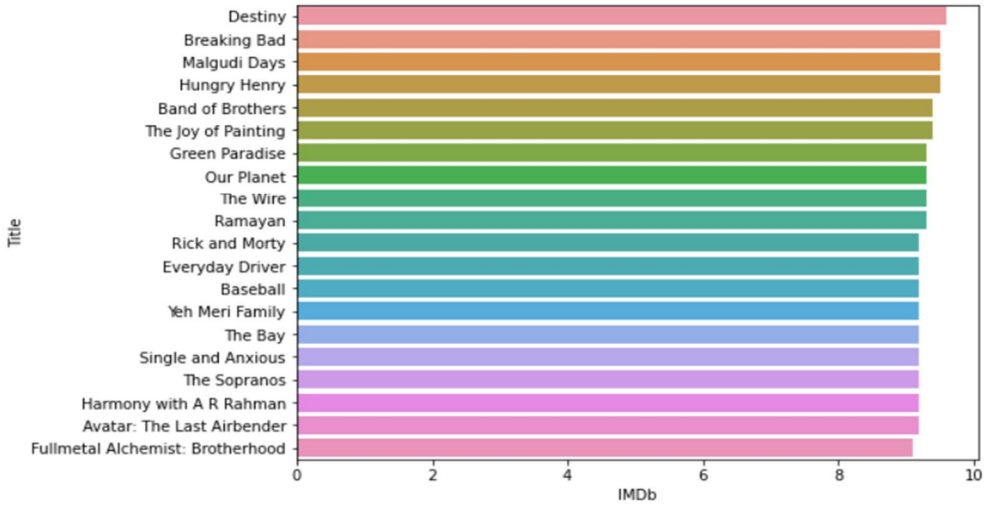


Fig 5: Data Visualization of top-rated TV shows

VI. RESULT DISCUSSION

Briefly illustrate the outcomes of the project development along with the benefits to society.

A. The proposed system should be able to predict the rating of various tv shows accurately.
B. The system should also be able to generate graphs based on various factors to gain meaningful insights into the popularity of TV shows.
C. The system will be able to plot graphs based on gender, age and location.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XII Dec 2021- Available at www.ijraset.com

VII. CONCLUSION

This is the conclusion that the project we have undertaken has been carried out with sincere effort. Most of the requirements are met up to the mark and the remaining requirements can be completed with a short extension. Designed to speculate model to predict the popularity of tv games based on IMDB databases. The model uses a database to predict estimates based on a variety of factors. The results will be accurate based on the data we will receive using data mining and machine learning. This project will cater to all the needs of students and will be of benefit to aspiring writers and beginners.

VIII. ACKNOWLEDGEMENT

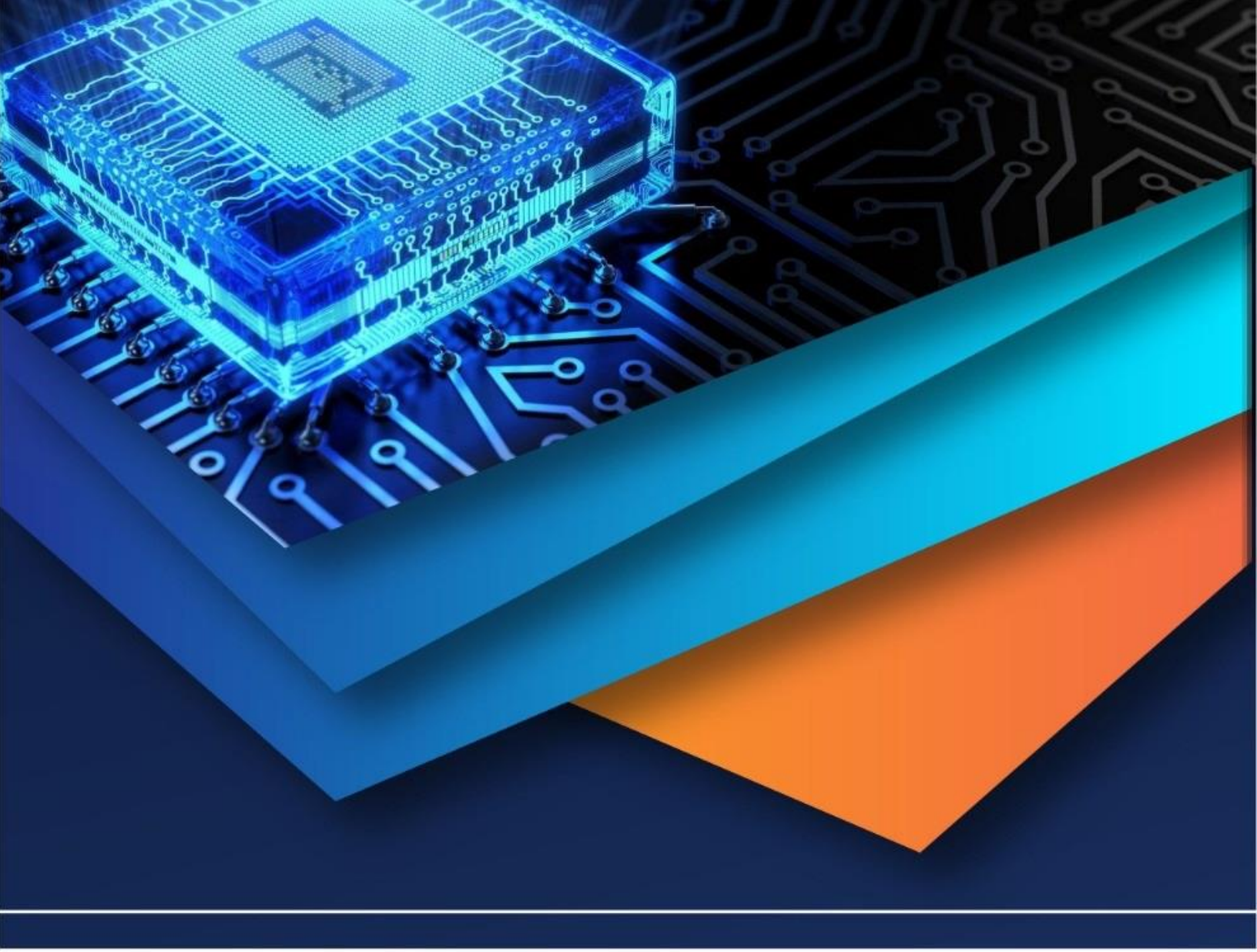
We give our sincere gratitude, a deep sense of approbation and respect to our guide and mentor Prof. Ajay Khatri, Professor, AITR, Indore for his motivation, sagacious guidance, constant encouragement, vigilant supervision and valuable critical appreciation throughout this project work, which helped us to complete the project on time.

We express profound gratitude and heartfelt thanks to Prof. Ajay Khatri, Professor CSE, AITR Indore for his support, suggestion and inspiration for carrying out this project. I am grateful to other faculty and staff members of the CSE Dept, AITR Indore for providing me with all support, help and advice during the project. We would be failing in our duty if we do not acknowledge the support and guidance received from Dr SC Sharma, Director, AITR, Indore whenever needed. We take the opportunity to convey my regards to the management of Acropolis Institute, Indore for extending academic and administrative support and providing me with all the necessary facilities for the project to achieve our objectives.

REFERENCES

[1] TV Show Analysis: IEEE
[2] TV Show Analysis using Data Mining
[3] Popularity analysis using Data Mining: IEEE
[4] http://202.62.95.70:8080/jspui/bitstream/123456789/13576/1/1NH16CS006%20030%20147TV%20Show%20Popularity%20Analysis%20using%20Data%20Mining.pdf

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)