



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 4, April 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com



Big Data Analytics for Program Popularity Prediction in Broadcast TV Industry

R S Wawre¹, Sandeep N Gite²

P.G. Student, Department of Computer Engineering, MSS's Engineering College, Jalna, Maharashtra, India¹

Assistant Professor, Department of Computer Engineering, MSS's Engineering College, Jalna, Maharashtra, India²

ABSTRACT: Television group of onlookers rating is a vital pointer as to prevalence of projects and it is likewise a factor to impact the income of communicate stations through promotions. Albeit higher evaluations for a given program are gainful for the two supporters and promoters, little is thought about the components that make programs increasingly alluring to watchers. So as to think about the prevalence of performers, we consider the quantity of hits gotten by the tweets identified with them on Twitter. In this project we are using three different data mining techniques namely – Decision Tree, Naïve Bayes, and XGBoost. The purpose of this work is to evaluate the performance of TV Show and also calculate how many people are liked to a particular show or actors of that show and predicting Popularity of that shows, based on the text reviews. We are comparing each data model with other techniques so that we get the most accurate results. The overall objective of our work is to predict more accurately, which TV show will gain more popularity in the future.

KEYWORDS: Dynamic time warping, random forests regression, decision tree, naïve bayes.

I. INTRODUCTION

Twitter has been a source of activity not just for users but also for third-party interest groups. In this context, there are increasing number of applications on Twitter which enhance its usability. An example is a user-friendly application that lets people streamline tweets from their mobile (example: Tweet Deck). The purpose of this project is to find out trends by aggregating the data in social networking site such as Twitter. Analysis of Twitter Data to evaluate sentiment to the user tweets, find the hash-tag used by the users and to find the most popular user for the twitter data related to Commonwealth Games Eid Festival India Vs England Test Match Gaza attack Launch of Xiaomi Smartphone on Flipkart Every day.

In this context, it is of great importance to predict the popularity of broadcast TV programs. First, using the program popularity prediction results, the audience will save much time when trying to find valuable TV programs among massive collections of video resources, which will improve user satisfaction and retention. Second, based on program popularity forecasting data, a company will be able to maximize its advertising effect by choosing the TV programs with highest potential. Finally, with the help of the popularity prediction model, a broadcast TV operator will be able to optimize the configuration of the network in advance by deploying enough transmission and storage resources to distribute popular programs.

We will be working on creating an algorithm which will decide whether a particular show will remain popular in the future or not, based on the comments we receive from Twitter or other blogging sites. We also have a tendency to develop a GUI that will assist the user and advertisers to evaluate a particular show. The success of the model will be predicted by using machine learning models. Analysis of the popularity of these TV shows will help both the audience and the corporate advertisers. Audience will be able to make choices based on the suggestions about what show to watch next and the corporate advertisers will be able to sell ads based on the popularity of these shows.

II. RELATED WORK

Due to the rich variety and timeliness of TV content, the semantic understanding of broadcast TV programs is more difficult than that of news, microblogging or other web content. An ideal prediction model for broadcast TV programs achieves not only high prediction accuracy but also good calculation performance, which means that the prediction result is available before audience interest fades. At present, there is little research on program popularity prediction for broadcast TV. The existing popularity prediction methods are for other media formats but



can be used as references. Commonly used web content popularity prediction methods include cumulative growth, temporal analysis and evolutionary trends.

1. The development of Facebook and other social media websites encouraged more people to share their thoughts. Undertakings saw this and began their very own fan pages for the clients to connect with them and express their steadfastness. Yu-Hsuan Chang, Chen-Ming Wu, Tsun Ku and GwoDong Chen taken a shot at an anticipating model(2013) that utilized the substance created in the television program fan pages by watchers and Artificial Neural Network to perform conjectures on program evaluations. The exploration developed a program evaluations estimate module dependent on Back-engendering Network. It at that point utilizes trained Artificial Neural Network to play out an appraisals conjecture for up and coming projects.

2. In 2017, Chengang Zhu, Guang Cheng and Kun Wang with Big Data to foresee the prevalence of communicate television programs. With the assistance of the forecast model, a communicate television administrator will most likely upgrade the setup of the system ahead of time by conveying enough transmission and capacity assets to disperse well known projects. Auto Regressive and Moving Average (ARMA) models were connected to genuine follows extricated from YouTube and exact forecasts were acquired. They connected a dynamic time warping(DTW) remove based k-medoids calculation to group programs with comparable fame into four transformative patterns, which can catch the innate heterogeneity of program ubiquity.

3. The social network service, is a rich source of information on customer. Examples include Facebook, MySpace, Dig, Twitter and JISC list serves on the academic side, because of its ease of use, speed and reach, social media is fast changing the public chat in society and setting trends and agendas in topics that range from the environment and politics to technology and the entertainment industry. Since social media can also be understood as a form of collective wisdom, it can be used to predict real-world outcomes. Twitter has millions of users that spread millions of personal posts on a daily basis. And this gives us the opportunity to study social human subjectivity. Manual classification of thousands of posts for opinion mining task is unfeasible for a human being.

4. Yusuke Fukushima, Toshihiko Yamasaki and Kiyoharu Aizawa of the University of Tokyo in 2016 taken a shot at an examination to gauge the appraisals of shows before their communicate without utilizing video substance or group of onlookers' reaction to them. They considered the ubiquity of performing artists included, every day perspectives onscreen characters' Wikipedia and notices of them on Twitter were utilized as markers of prevalence. It was clear that a show gets a higher rating if various famous on-screen characters are in it. Anticipate the gathering of people evaluations with high exactness by considering the performers and staff associated with expansion to the traditional data, for example, broadcast appointments and the stations in which they play.

III. METHODOLOGY

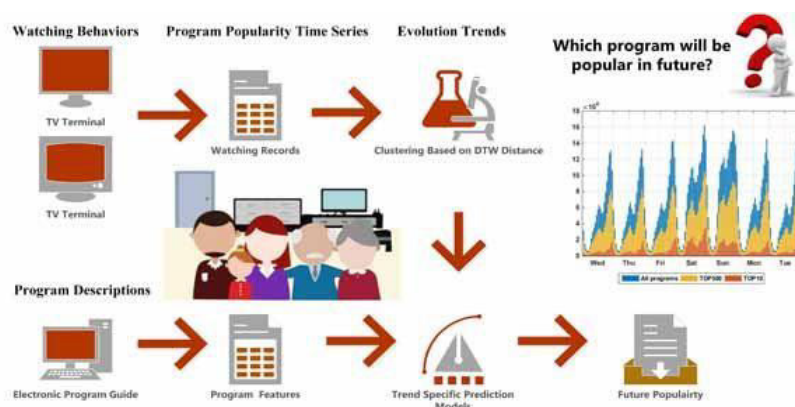
A. **PROBLEM STATEMENT**-The program popularity prediction problem can be defined as follows. Let $c \in C$ be an individual program from a set of programs C that are observed during a period T . We use $t \in T$ to describe the age of a program (i.e. the time since it was first published) and mark two important moments: the indication time t_i , which is the time at which we perform the prediction, and the reference time t_r , which is the moment of time for which we want to predict program popularity. Let $N_c(t_i)$ be the popularity of c from the time a program was published until t_i and $N_c(t_r)$ be the value that we want to predict, i.e., the popularity at a later time $N_c(t_r)$. We define $N_c(t_i, t_r)$ as the prediction outcome: the predicted popularity of program c at time t_r using the information available until t_i . Thus, the better the prediction, the closer $N_c(t_i, t_r)$ is to $N_c(t_r)$.

A predictive model is required to be built which works as a rating system that'll be useful for people who are willing to watch a particular show. They will also be able to get feedback from previous viewers and provide help to new viewers. Vast databases from social media sites will be used for the same.

B. STEPS-

1. We search for a particular keyword which is regarding a TV series or Web Series or Movies and then extract the tweets related to Keyword.
2. To ensure that we obtained all tweets referring to a TV show, we used hash tags of TV show as search arguments. Tweets were extracted over frequent intervals using the Twitter Search API.
3. Twitter users can simply create a hash tag by prefixing a word with a hash symbol "#hashtag." For instance, various hashtags, such as #balikavadhu, #BALH, #CID had been created for different TV shows.
4. Clustering analysis allows an object not to be part of a cluster, or strictly belong to it, calling this type of grouping hard partitioning. On the other hand, soft partitioning states that every object belongs to a cluster in a determined degree.

5. The analysis of Sentimental comment and predicting whether it is good or bad comments. Easy prediction of TV Show trending based on people rating. Good or Bad comments based on peoples reviews or comments. Easy importing of data and exporting it into the graph. Graphical data in the printable format. The visitor will get to know the show popularity.



IV. CONCLUSION

In this paper, we have analysed massive user behaviour data and presented our improved method to predict the popularity of broadcast TV programs. To the best of our knowledge, this is the first work to tackle the problem of predicting program popularity in the broadcast TV platform. We have presented a predictive model to predict the popularity of TV shows based on user comments from social media. We are able to obtain significant results over the provided datasets. The model uses sentiments of the viewers and can be implemented for any genre of TV show. Results are highly accurate based on the values we've obtained using data mining and machine learning.

REFERENCES

- [1] Yu-Hsuan Cheng, Chen-Ming Wu, Tsun Ku, Gwo-Dong Chen. A Predicting Model of TV Audience Rating Based on Facebook, 2013.
- [2] Yusuke Fukushima, Toshihiko Yamasaki, Kiyoharu Aizawa. Audience Ratings Prediction of TV Dramas Based on the Cast and their Popularity, 2016.
- [3] Nicolai H. Egebjerg, Niklas Hedegaard, Gerda Kuum, Raghava Rao M, Ravi Vatrappu. Big Social Data Analytics in Football: Predicting Spectators and TV Ratings from Facebook Data, 2017.
- [4] Chengang Zhu, Guang Cheng, Kun Wang. Big Data Analytics for Program Popularity Prediction in Broadcast TV Industries, 2017.
- [5] Javaria Ahmed, Prakash Duraisamy, Amr Yousef, Bill Buckles. Movie Success Prediction using Data Mining, 2017. S. Bhuvaneswari, T. S. Subashini, "Automatic Detection and Inpainting of Text Images", International Journal of Computer Applications (0975 – 8887) Volume 61– No.7, 2013
- [6] Daniel Gruhl, R. Guha, Ravi Kumar, Jasmine Novak and Andrew Tomkins. The predictive power of online chatter. SIGKDD Conference on Knowledge Discovery and Data Mining, 2005.
- [7] G. Mishne and N. Glance. Predicting movie sales from blogger sentiment. In AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs, 2006.
- [8] Mahesh Joshi, Dipanjan Das, Kevin Gimpel and Noah A. Smith. Movie Reviews and Revenues: An Experiment in Text Regression NAACL-HLT, 2010.
- [9] Ramesh Sharda and Dursun Delen. Predicting box office success of motion pictures with neural networks. Expert Systems with Applications, vol 30, pp 243– 254, 2006.
- [10] Sitaram Asur, Bernardo A. Huberman "Predicting the Future With Social Media" 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology.
- [11] N. B. Hassine, R. Milocco, and P. Minet, "ARMA based popularity prediction for caching in content delivery networks," in Proc. Wireless Days, Mar. 2017, pp. 113–120.



- [12] K. Wang, J. Mi, C. Xu, Q. Zhu, L. Shu, and D.-J. Deng, "Real-time load reduction in multimedia big data for mobile Internet," ACM Trans. Multimedia Comput. Commun. Appl., vol. 12, no. 5s, p. 76, 2016.
- [13] F. Figueiredo, J. M. Almeida, M. A. Gonçalves, and F. Benevenuto, "TrendLearner: Early prediction of popularity trends of user generated content," Inf. Sci., vols. 349–350, pp. 172–187, Jul. 2016.
- [14] M. Tsagkias, W. Weerkamp, and M. de Rijke, "News comments: Exploring, modeling, and online prediction," in Advances in Information Retrieval. Cham, Switzerland: Springer, 2010, pp. 191–203.
- [15] A. Tatar, P. Antoniadis, M. D. de Amorim, and S. Fdida, "From popularity prediction to ranking online news," Social Netw. Anal. Mining, vol. 4, pp. 174–183, Dec. 2014.



**Impact Factor:
7.512**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com



www.ijirset.com

Scan to save the contact details