

To Evaluate & Predict the Television Serials' TRP

^[1] Amritesh Sidhu, ^[2] Er. Kanwal Preet S. Attwal, ^[3] Dr. Gaurav Gupta

^[1] Research Scholar, ^{[2][3]} Assistant Professor

^{[1][2][3]} Department of Computer Science and Engineering, Punjabi University, Patiala

^[1] guptasabia@yahoo.com, ^[2] kanwalp78@yahoo.com, ^[3] gaurav.shakti@gmail.com

Abstract: - In present scenario, the number of TV reality shows are growing day-by-day. Television rating points (TRPs) play a vital role in the broadcasting field of TV. TRPs can also be deployed to study the interests of the viewers. In this paper, firstly viewers' opinions are examined. Then six factors have been considered based on which, a random tree is made. This random tree is a part of classification (a technique of data mining). Finally, this random tree helps to predict the TRP of a serial or a channel.

Keywords: - Popularity prediction, Sentiment analysis, Data mining, Opinion mining, TV Rating, WEKA, Classification.

1. INTRODUCTION

1.1 Overview

Data mining is defined as extracting the information from huge set of data. In other words data mining is the procedure of mining knowledge from data. Data mining also involves other processes such as data cleaning, data integration, data transformation, pattern evaluation and data representation.

1.2 Data mining and knowledge discovery databases

Data mining, likewise known as knowledge discovery in databases (KDD), refers to the nontrivial extraction of verifiable, already obscure and possibly helpful data from information in databases. Knowledge discovery in databases (KDD) is the process of discovering useful knowledge from a collection of data. This data mining technique is a process that includes data preparation and selection, data cleansing, incorporating prior knowledge on data sets and interpreting accurate solutions from the observed results [11]. The accompanying (Figure 1.1) indicates data mining as a stage in an iterative learning revelation process.

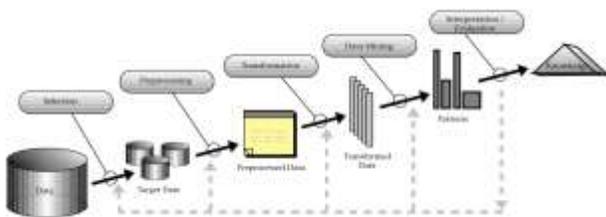


Figure 1.1: KDD Process [11]

The iterative procedure comprises of the accompanying advances:

- **Data cleaning:** It is a stage in which information and super information are expelled from the gathering.
- **Data coordination:** At this stage, numerous information sources, regularly heterogeneous, might be joined in a typical source.
- **Data selection:** At every one of the information gathered by the client is not completely required. Here pick that information which is considered as useful for data mining [11].
- **Data change:** It is a stage in which the information is changed into shapes suitable for the mining strategy.
- **Data mining:** It is the process that includes searching for patterns or a set of representations as classification trees, clustering or regression.

1.3 Overview of television serials TRP

Today TV has turned into a basic part of our lives. Would you be able to imagine a world without TV? No serials, news channels or cricket matches. Will it not be an exhausting world? Is it safe to say that it isn't awesome that a newsreader that is several kilometers away in a news studio can be seen and heard by simply exchanging on your TV? In any case, TV is a current invention [4]. TV step by step developed as a medium of following two decades.

From being "radio with pictures" it obtained a one of its very own kind style. When television started in India 'Doordarshan' was the sole provided for the TV services. Then DD2 later on DD Metro were another channels by the government. After that private and foreign broadcasters and cable television came into existence [18].

Unscripted TV is another new pattern. Have you seen

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 6, Issue 9, August 2019

programs such as 'Indian idol', 'big boss' and 'sa re ga ma pa'? They are cases of reality TV. Along these lines, in this time of web, TV has been continually reevaluating itself as a medium. The early transmissions were all highly contrasting. The primary effective program in shading was transmitted by Columbia Broadcasting System (CBS) in USA in 1953 [12].

For the greater part of the twentieth century, the main approaches to watch television were through finished the air communicates and link signals. With communicate TV, a receiving wire gets radio waves to transmit pictures and sound to your TV. With satellite TV, wires associate with a set-top box or to your TV itself [4].

These wires keep running from your home to the closest satellite TV station, which goes about as only major receiving wire. Beside a couple of choices like satellite TV, communicate and link are the fundamental approaches to sit in front of the TV [7].

Relationships between audience behavior and TV program content can be identified by integrating TV ratings with multimedia content. Understanding audience behavior is useful for knowing the interest of people. The main aim is to prepare a framework for mining TV audience behavior from TV ratings [16].

An important relationship between the percentage of negative and positive views and the audience of an episode of a TV series is suggested. Networks could identify strengths and weaknesses of TV shows earlier to maximize course-correct marketing activities [9].

Web TV, is video and sound conveyed over an Internet association. It is otherwise called Internet convention TV, or IPTV (internet protocol television). You can watch Internet TV on a PC screen, a TV screen (through a set-top box) or a cell phone like a mobile phone or an iPod [7].

The purpose of this work is to evaluate the performance of TV show and also calculate the likes of a particular show or actors of that show and predicting popularity of that shows based on the text reviews. The reviews are getting on social networking websites like Facebook. [20].

It's nearly the same as getting TV through a receiving wire or a progression of link wires the distinction is that data is sent over the Internet as information. In the meantime, you can discover significantly more assortment on Internet TV than digital TV. Web TV is generally new there are loads of various approaches to get it, and quality, substance and expenses can differ significantly. Conventional TV systems are likewise slipping into the innovation and exploring different avenues regarding diverse organizations [1].

You can watch two fundamental sorts of communicates through Internet TV: live communicates or on-request recordings. In spite of the two essential communicate classes, there are three fundamental expense structures for Internet TV:

- **Free:** Aside from the expense you pay for Internet network, numerous Internet TV destinations or channels don't cost anything. A large number of these free locales are bolstered by publicizing, so standard promotions may appear around the site, or short ads may play before you watch recordings.
- **Subscription:** This works simply like your link charge. You ordinarily pay a month to month expense for a specific number of channels or on-request video.
- **Pay-per-see:** Pay-per-see recordings or podcasts can cost nothing if the webpage is free, and real systems by and large charge amongst \$3 and \$7 for downloads and rentals.

2. METHODOLOGY

2.1 Pre requisite

The indispensable sources of data and gadgets for data gathering are the entire system. These are some ways:

- ✓ Viewers Opinions
- ✓ Data mining devices
- ✓ Opinion records
- ✓ Programming language

2.2 Techniques used

1. Classification: Classification is a data mining technique that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data. For example in case of TRP we can make decision trees or classes that how many people like or dislike TV serials. The learning and classification steps of a decision tree are simple and fast. Classification process includes two steps:

- **Building the model:** In this the classifier is built from the training set made up of database tuples and their associated class labels.
 - **Using classifier for classification:** The accuracy of classification rule is estimated by the test data. The classifier is used for classification.
- 2. Prediction:** Prediction in data mining is to identify data points purely on the description of another related data value. Prediction derives the relationship between a thing you know and a thing you need to predict for future reference. For numerical prediction the statistical methodology is used for regression analysis. For example in case of TRP what will be the views of

people in future for any TV serial.

- Robustness is the ability of predictor to make correct predictions from noisy data.
- Scalability refers to construct the predictor efficiently.
- Accuracy of the predictor refers to how a predictor can guess the value of predictor attribute for a new data.

2.3 Algorithm

1. Finding the main factors.
2. Determine the influencing factors.
3. Analyse the viewer’s satisfaction.
4. Impact of TRP productive on viewer’s satisfaction.
5. Investigate the gathered rating based reaction by classification tree.
6. Predicting a viewer’s opinion. Figure 4.2 shows the functionality of TRP.

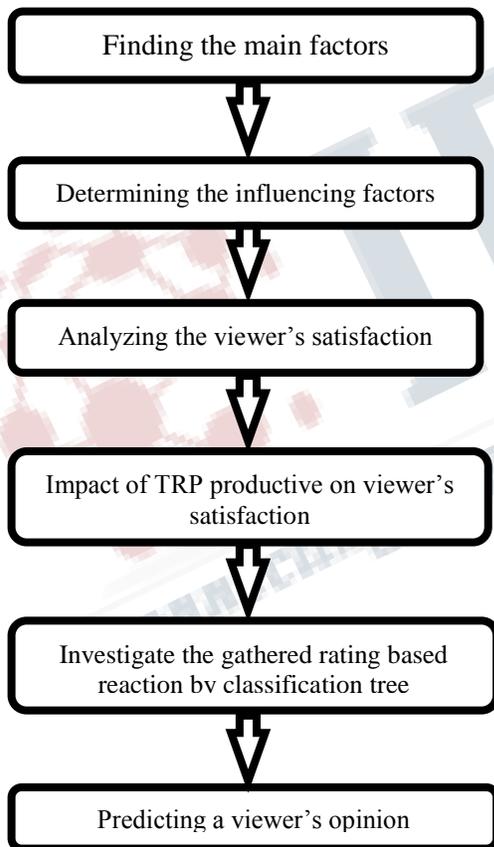


Figure.2.3.1: Functionality of TRP

2.4 Data set

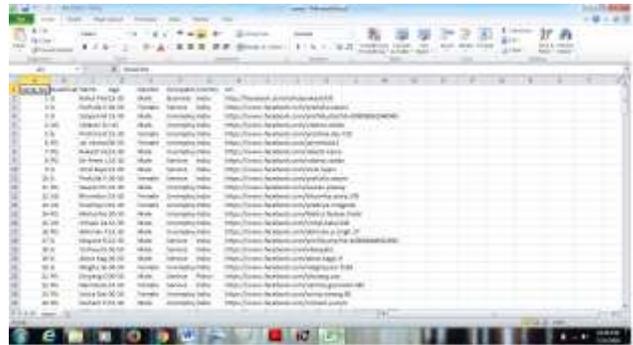


Figure 2.4.1: Historical data set

Figure 2.4.1 shows the historical data set based on the Facebook. Used for prediction of television rating points.

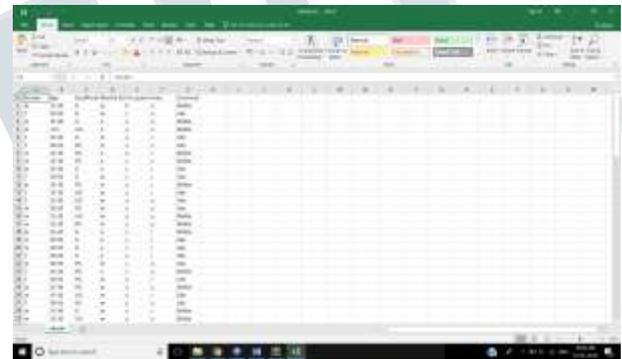


Figure 2.4.2: Predicted data set

Figure 2.4.2 shows the predicted data set. Prediction has been created on the basis of historical data. We considered 6 Major factors for prediction which were:-

- Gender
- Age
- Qualification
- Marital status
- Occupation
- Area

On these factors we created condition which are like and dislike.

3. RESULTS AND DISCUSSION

A process is followed and experimental results are designed to show the importance of television rating points in the field of television and film industry. We first examine the opinion of viewers on the basis of historical data. A prediction is created on the basis of historical data. In this process six genuine factors are examined. The data

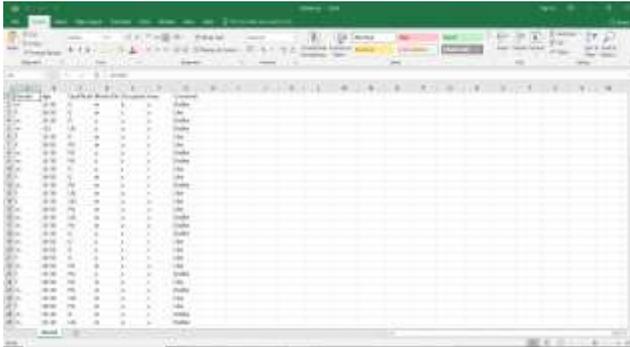
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 6, Issue 9, August 2019

related to these factors has been taken from facebook. Prediction will be created on the basis of historical Data. We considered 6 major factors for prediction which were Gender, Age, Qualification, Marital Status, Occupation and Area. Based on these factors a random tree is made. Random tree which is a part of classification is predicting the TRP of a channel or serial.

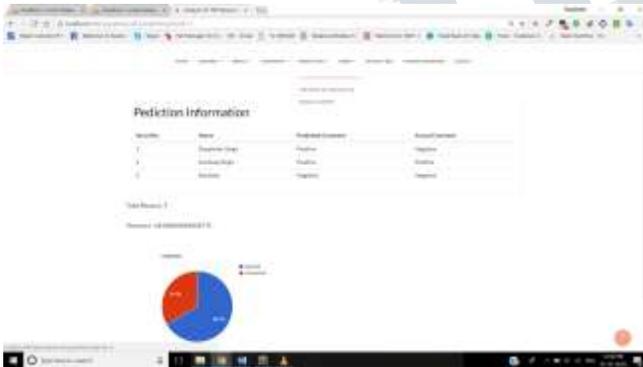
Analysis Of TRP (Television Rating Points)

Data Collection:



Channel	TRP	Channel	TRP
1	100	1	100
2	100	2	100
3	100	3	100
4	100	4	100
5	100	5	100
6	100	6	100
7	100	7	100
8	100	8	100
9	100	9	100
10	100	10	100
11	100	11	100
12	100	12	100
13	100	13	100
14	100	14	100
15	100	15	100
16	100	16	100
17	100	17	100
18	100	18	100
19	100	19	100
20	100	20	100
21	100	21	100
22	100	22	100
23	100	23	100
24	100	24	100
25	100	25	100
26	100	26	100
27	100	27	100
28	100	28	100
29	100	29	100
30	100	30	100
31	100	31	100
32	100	32	100
33	100	33	100
34	100	34	100
35	100	35	100
36	100	36	100
37	100	37	100
38	100	38	100
39	100	39	100
40	100	40	100
41	100	41	100
42	100	42	100
43	100	43	100
44	100	44	100
45	100	45	100
46	100	46	100
47	100	47	100
48	100	48	100
49	100	49	100
50	100	50	100

Prediction: You can get prediction of a particular Channel.



Summary

In this research work, views of people have been taken from historical data. For getting the information of people, six factors have been created on the basis of which random tree is made. Random tree which is a part of classification is predicting the TRP of a channel or serial.

4. CONCLUSION AND FUTURE SCOPE

In this paper, the historical data has been analyzed to obtain trends in the viewership. By the help of historical data the future of the television serials are predicted. The classification tree helps to classify the historical demographic data & obtain the predicted future viewers.

The WEKA tool was used to analyze the whole process to ensure the viewer's choice for future. The programming language PHP was used to implement the webpage for TRP representation.

- ✓ Prediction will be created on the basis of historical data so the real time data can be used for the prediction.
- ✓ In this 6 major factors were considered so more factors can be added.
- ✓ On behalf of these factors like & dislike conditions were created, in future comments can also be applied.
- ✓ Random tree was created on the basis of historical data, the other trees can be applied for it.
- ✓ WEKA tool was used for classification; In future other data mining tools can also be tried.

REFERENCES

- [1]. A. Bria, P. Karrberg, P. Andersson (2007), "TV in the mobile or TV for the mobile: challenges and changing value chains", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1-5.
- [2]. A. Mhaisgawali, N. Giri (2014), "Detailed descriptive and predictive analytics with twitter based TV ratings", International Journal of Computing and Technology, 1(4), pp. 125-130.
- [3]. A. Rashid, N. Anwer, M. Iqbal, M. Sher (2013), "A survey paper: areas, techniques and challenges of opinion mining", International Journal of Computer Science Issues, 10(2), pp. 18-31.
- [4]. A. Singh, S. K. Mehta, H. G. Mishra (2011), "TRP as a measure of visual communication: a study of Jammu city, India", International Conference on Economics and Finance Research, 2(7), pp. 123-137.
- [5]. B. M. Ramageri (2014), "Data mining techniques and applications", Indian Journal of Computer Science and Engineering, 1(4), pp. 301-305.
- [6]. D. Anand, A. V. Satyavani, B. Raveena, M. Poojitha (2018), "Analysis and prediction of television show popularity rating using incremental k-means algorithm", International Journal of Mechanical Engineering and Technology, 9(1), pp. 482-489.
- [7]. E. Panova, A. Raikov, O. Smirnova (2015), "Cognitive television viewer rating", International Young Scientists Conference on Computational Science, 66, pp. 328-335.
- [8]. G. Vinodhini, R. M. Chandrasekaran (2012), "Sentiment analysis and opinion mining: a survey", International Journal of Advanced Research in

**International Journal of Engineering Research in Computer Science and Engineering
(IJERCSE)****Vol 6, Issue 9, August 2019**

- Computer Science and Software Engineering, 2(6), pp. 283-292.
- [9]. L. Molteni, J. P. D. Leon (2016), "Forecasting with twitter data: an application to USA TV series audience", International Journal of Design and Nature and Ecodynamics, 11(3), pp. 220-229.
- [10]. M. J. Panaggio, P. W. Fok, G. S. Bhatt, S. Burhoe, M.Capps, C. J. Edholm, F. E. Moustaid, T. Emerson, S. L. Estock, N. Gold, R. Halabi, M. Houser, P. R. Kramer, H. W. Lee, Q. Li, W. Li, D. Lu, Y. Wian, L. F. Rossi, D. Shutt, V. C. Yang, Y. Zhou (2016), "Prediction and optimal scheduling of advertisements in linear television", Mathematics Department, Rose-Hulman Institute of Technology, pp. 2-24.
- [11]. N. Rikhi (2015), "Data mining and knowledge discovery in database", International Journal of Engineering and Trends Technology, 23(2), pp. 64-70.
- [12]. P. G. Malur, N.Lakshmikantha, V.Prashanth(2014),"Reeling the reality: A study on contemporary reality shows and their Influence on other entertainment program genres", International Research Journal of Social Sciences,3(8), pp. 35-38.
- [13]. P. Jain, P. Jakate, A. Dhotre, J.Bhati (2015), "A novel approach to analysis of TV shows using social media, machine learning and big data", International Journal of Technological Exploration and Learning, 4(6), pp. 604-612.
- [14]. R. N. Rao, H. Vani, S. Vandana (2015),"A study of viewers satisfaction towards hindi news channels at Hyderabad", Indian Journal of Commerce and Management Studies, 6(1), pp. 62-69.
- [15]. R. Pagano, M. Quadrana, P. Cremonesi, S. Bittanti, S. Formwentin, A. Mosconi (2015), "Prediction of TV ratings with dynamic model", ACM Workshop on Recommendation System for Television and Online Video,
- [16]. R. Hinami, S. Satoh (2017), "Audience behavior mining: integrating TV ratings with multimedia content", IEEE Computer Society, 24(2), pp. 44-54.
- [17]. S. Pankanti, S. Chavan, M. Kumar, P. Vitkar (2017), "A TV show suggestion framework In view of Viewer's rating", International Journal of Engineering Research and Technology, 10(1), pp. 370-373.
- [18]. S. Patil, Y. Jog, S. Pareek, D. Chodnekar (2015), "Challenges and opportunities of television rating point (TRP) and television audience measurement (TAM) in india", International Journal of Advanced Technology in Engineering and Science, 3(10), pp. 23-34.
- [19]. S. Sereeday, J. Cui (2017),"Using machine learning to predict future TV ratings", Nielsen Journal of Measurement, 1(3), pp. 3-12.
- [20]. T. Kadam, G. Saraf, V. Dewadkar, P. J. Chate (2017), "TV show popularity prediction using sentimental analysis in social network", International Research Journal of Engineering and Technology, 4(11), pp. 1087-1089.