

TV Show Popularity Prediction using Sentiment Analysis in Social Network

Tejaswi Kadam¹, Gaurav Saraf², Vikas Dewadkar³, P.J Chate⁴

^{1,2,3}Student, Computer Department, Bharati Vidyapeeth's College of Engineering Lavale, Pune, Maharashtra

⁴Professor, Computer Department, Bharati Vidyapeeth's College of Engineering Lavale, Pune, Maharashtra

Abstract - TV Show Popularity prediction using sentiment analysis is one of the most interesting and challenging tasks. A critical demand along this line is to predict the popularity of online serials, which can enable a wide range of applications, such as online advertising, and serial recommendation. The problem motivation stated above suggest is that it is only the viewer of a program who is responsible for its popularity or failure and if we anyhow can identify the most common features of a program which, the viewers want most, and through some effective scientific methodology could insert these requirements in the proposed TV program well at the time of production.

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 getting reviews on social networking websites like Twitter.

Key Words: Popularity prediction, Sentiment Analysis, Natural Language Processing, Porter Stemmer

1. INTRODUCTION

With the rapid development of sharing Websites, more and more people would like to become audiences in their daily entertainments. such as TV shows and Web sites, to attract more audiences. although many efforts have been taken for the popularity prediction. Also, episodes released on weekends or holidays may attract more audiences than those on workdays. Furthermore, since different episodes are usually released on different days, Therefore, the popularity prediction for tv shows is one of the most interesting and challenging tasks. 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. Reality TV is the new mantra of television producers and channel executives. the Main purpose is to TRP ratings. Nowadays Most of the television shows are reality shows specializing in dancing, singing, and acting. We conclude to build such a system that will recognize people's sentimental comments on TV shows. The tweets related to the particular show will be extracted. The comments will be

gathered from various sources social networking websites like Twitter. On the basis of people's comment and the TV Show popularity will be rated accordingly. The system allows admin to add text views likes-dislikes and their sentimental comment. Based on the people's comment, Visitors can view TV show popularity data Visitor can also view the popular show rating as well as the top show in a country.

2. RELATED WORK

With the rapid development of sharing Websites, more and more people would like to become audiences in their daily entertainments. such as TV shows and Webisodes, to attract more audiences. although many efforts have been devoted to the popularity prediction. Also, episodes released on weekends or holidays may attract more audiences than those on workdays. Furthermore, since different episodes are usually released on different days, Therefore, the popularity prediction for TV shows is one of the most interesting and challenging tasks. 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. Reality TV is the new mantra of television producers and channel executives. the Main purpose is to TRP ratings. Nowadays Most of the television shows are reality shows specializing in dancing, singing, and acting. We conclude to build such a system that will interpret people's sentimental comments on TV shows. The tweets related to the particular show will be extracted. The comments will be gathered from various sources social networking websites like Twitter. On the basis of comments given by the people and sentiments, the TV Show popularity will be rated. The system allows admin to add text views likes-dislikes and their sentimental comment. Based on the people's comment, a Visitors can view TV show popularity data Visitor can also view the popular show rating as well as the top show in a country.

3. PROPOSED SYSTEM

The figure shows the architecture of the proposed system. The system flow is divided into followings way. Firstly the system fetches TV show related data from twitter API where text reviews related to the particular show, actor, and director are given..

A. Hash tagged data set

To create the hashtagged dataset, we first filter out duplicate tweets, non-English tweets, and tweets that do not contain hashtags. From the remaining set (about 4 million), we investigate the distribution of hashtags and identify what we hope will be sets of frequent hashtags that are indicative of positive, negative and neutral messages. These hashtags are used to select the tweets will be used for development and training

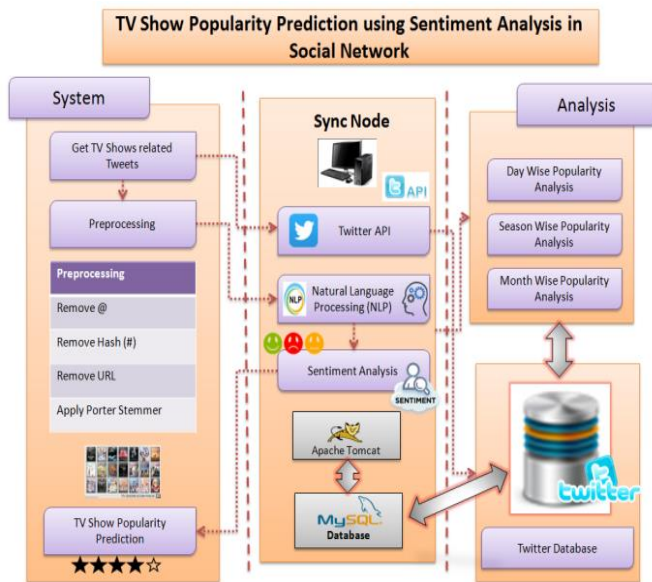


Fig-1: Architecture of TV Show Popularity Prediction using Sentiment Analysis in Social Network

Pre-processing is one of the important steps in text mining, Natural Language Processing (NLP) and information retrieval (IR). which gives tokenization, normalization i.e. remove @,remove #and URL. Data pre-processing is used to extract interesting and non-trivial knowledge from unstructured text data. Information Retrieval is important for deciding which documents in a collection should be retrieved so that we can satisfy a user's need for information.

B. Tokenization

The process of infringement a flow of text into words, symbols, phrases, or other meaningful elements called tokens. The list of tokens becomes input for the further

processing such as parsing or text mining. It splits sentences into words. Textual is only a block of characters at the beginning. All processes in order recovery require the words of the data set. Hence, the requirement for a parser is a tokenization of documents. This may sound slight as the text is already stored in machine-readable formats.

C. Normalization

To carry out processing on natural words manuscript, it is essential to perform normalization that mostly involves eliminating the punctuation, converting the entire text into lowercase or uppercase, converting numbers into words, expanding abbreviations, canonicalization of text, removes stop words from input text data. Stop words are the word that is automatically omitted from a computer-generated concordance or index.

D. Part-of-speech (POS) tagging

It detects if the word token is noun, verb, and adjective. POS Tagging in which a word is assigned in accordance with its syntactic functions.

E. Apply NLP:

One of the major challenges in natural language processing is teaching computers to understand the way humans learn and use language. Google search engine base their machine translation technology on NLP deep learning models. This model allows algorithms to read the text on a webpage, evaluate its meaning and translate it into another language. NLP algorithms are typically based on machine learning algorithms. as a substitute of manual coding large sets of rules, NLP can rely on machine learning to automatically learn these rules by analyzing a set of examples (i.e. a large corpus, like a book, down to a collection of sentences), and making a statically inference.

F. Sentiment Analysis

Sentiment analysis is another primary use case for NLP. Using sentiment analysis, data scientists can assess comments on social media to see how their business's brand is performing, for example, or review notes from customer service teams to identify areas where people want the business to perform better.

G. Analysis

Day Wise Popularity Analysis: We can analyzes one day popularity like live show is a one day show so user want to find particular date wise analysis.

Season wise Popularity Analysis: In three seasons which type of shows was to hit in given particular period this type analysis done in season wise analysis.

Month wise Popularity Analysis : We can also calculate month wise popular shows list and watch it. Like day and season wise analysis we analyze a month wise.

4 . PROPOSED ALGORITHM

PORTER STEMMER ALGORITHM

Is a process for removing the commoner morphological and in flexional endings from words in English. Following are the steps of this algorithm:-

Steps of Algorithm:

- Step1.** Gets clear of plurals and -ed or -ing suffixes
- Step2.** Turns terminal y to i when there is another vowel in the stem
- Step3** Maps double suffixes to single ones: -ization, -ational, etc.
- Step4.** Deals with suffixes, -full, -ness etc.
- Step5.** Takes off -ant, -ence, etc. Removes a final -e

5. CONCLUSIONS

Here we present a system TV Show Popularity Prediction using Sentiment Analysis in Social Network for users, which predict the popularity of the show among several shows, actors, reality shows and serials based on based on the text reviews which are getting from social networking websites like Twitter. The advantages of using this system are that it helps in analyzing TV Show details and helps to rate prediction based on Twitter tweets.

REFERENCES

- [1] " T. Jeon, 1. Cho, S. Lee, G. Baek, and S. Kim, "A movie rating prediction system of user propensity analysis based on collaborative filtering and fuzzy system," in Fuzzy Systems, 2009. FUZZ-IEEE 2009. IEEE International Conference on. IEEE, 2009, pp. 507-511.
- [2] O. Bora Fikir, İ lker O. Yaz, Tansel Özyer 2010 International Conference on Advances in Social Networks Analysis and Mining "A Movie Rating Prediction Algorithm with Collaborative Filtering" on. IEEE, 2009, pp. 507-511.
- [3] Jun Ai, Linzhi Li, Zhan Su, Chunxue Wu "Online-rating prediction based on an improved opinion spreading approach" 2017 29th Chinese Control And Decision Conference (CCDC)

- [4] G. Adomavicius and A. Tuzhilin, "Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", IEEE Transactions on Knowledge and Data Engineering 17 (2005), 634-749.