

# Evaluation of the Television Dramas Ranking Using the Bayes' Theorem

Bo-Yang LU<sup>1,a</sup>, Ji LI<sup>1,b</sup>, Yin-Zhu CHEN<sup>1,c</sup> and Hang XU<sup>1,d,\*</sup>

<sup>1</sup>Xuzhou Medical University, Xuzhou, Jiangsu, China

<sup>a</sup>2530485057@qq.com, <sup>b</sup>2539026957@qq.com, <sup>c</sup>1225100278@qq.com <sup>d</sup>chaoseu@163.com

\*Corresponding author

**Key words:** TV Drama, IMDB, Bayes' Theorem, Rank

**Abstract.** With the development of the TV industry, the prediction of the TV drama is becoming more and more important. In recent years, large data analysis has gradually developed, so through large data to analysis the trend of TV drama will be the greatest extent to accelerate the development of the TV industry. We referred to the IMDB, using Bayes' theorem, got the formula to calculate of the TV drama weighted score. Using the score to rank the TV dramas, reducing the impact of single factor gap. The top ten TV dramas are Battle of Changsha, All Quiet in Peking, The Great Protector, The Stand-In, Dating Hunter, Acacia flowers, Ordinary world, The Merchants of Qing Dynasty, The Chinese Old Peasant, HEY Daddy.

## Introduction

The increased demand for culture and entertainment after the affluent economy will drive and push the industry to grow rapidly and become one of the most fast-growing businesses in the future. With the development of the TV industry, the prediction of the TV drama is becoming more and more important. The current Chinese television market is highly competitive and full of different styles and topics. More than half of all TV productions each year go to waste without a buyer post-production, creating in a massive waste of resources. So the quality is the problem we need to solve. In recent years, large data analysis has gradually developed, so through large data to analysis the trend of TV drama will improve the TV drama quality and accelerate the development of the TV industry. In the paper, we refer to the IMDB, using Bayes' theorem to rank the TV dramas in order to improve the TV drama quality.

## Methods

### Data

Our data are the score in Appendix I: This is the average of the users' rating Evaluation information of TV play.

### Model Establishment and Analysis

In order to rank the TV drama, we used the well-known movie ranking website IMDB score, the Bayesian theorem of statistical methods [1-2]. By Bayesian inference require to be started from different assumptions of the initial degree of confidence, collecting new information, then according to the new information to adjust the degree of confidence. Adjusting the degree of confidence means closer to or even overthrowing the initial assumptions [3]. In this method, we will further score to get a more reliable score than the Appendix I.

According to the IMDB scoring method, the Bayesian theorem, Obtained a way of calculating the weighted score of each drama. The way to get the weighted score is in accordance with a reasonable proportion to calculate the fraction of data for different gravities. Bayes' theorem is a process that combined prior probability and conditional probability to derive the posterior probability.

The formula is

$$\text{weighted rank}(WR) = (v \div (v + m)) \times R + (m \div (v + m)) \times C \quad (1)$$

#### Explanation

$R$  = average for the movie (mean) = (Rating) (this average score is calculated by the average method)

$v$  = number of votes for the movie = (votes) (The number of voters will be counted only by frequent voters)

$m$  = minimum votes required to be listed in the top 250 (currently 1250) (The minimum number of votes required to enter the IMDB top 250)

$C$  = the mean vote across the whole report (currently 6.9) (The current average score for all movies)

#### Model Solving

Processing the data in the attachment, Obtain each value calculated by the weighted score formula using the Excel.

The following is the treatment of each item in the Appendix I.

$R$  stands for the score given in attachment 2

$v$  is the number of senior fans of the drama, use segmentation function to establish the number of senior fans who included in different votes.

Because the overall number of comments in the Appendix I is small, According to the IMDB standard,  $m$  should be 1250, Generally the highest number of reviews is 50,000 to 80,000, but the data only 26,000, so  $m$  is set to 700.

$C$  stand for the average of all TV dramas which have comments. After treatment, the top 20 TV dramas are shown in the table

Table Top 20 TV dramas

Name	score
Battle of Changsha	9.049268653
All Quiet in Peking	8.655918192
The Great Protector	8.115235852
The Stand-In	8.069594161
Dating Hunter	8.036112161
Acacia flowers	7.790175149
Ordinary world	7.694088201
The Merchants of Qing Dynasty	7.680898338
The Chinese Old Peasant	7.643728867
Hey Daddy	7.47312032
Red Sorghum	7.413455165
Honey Bee Man	7.30564112
The City of Warriors	7.235874305
Xiaoping Deng	7.187339311
GROW UP	7.176906673
Divorce Lawyers	7.165693663
One servant two masters	7.13848177
National Audit	7.134395311
Revelation of Life	7.100465272
Obstetrician	7.008166966

## Conclusions

We use the weighted score to rank, we prevented the impact of single factor gap is too large, get the top ten TV series. The top 10 TV dramas are Battle of Changsha, All Quiet in Peking, The Great Protector, The Stand-In, Dating Hunter, Acacia flowers, Ordinary world, The Merchants of Qing Dynasty, The Chinese Old Peasant, Hey Daddy.

The model built in this topic can study the score and development trend of the TV series, and the TV production team has a greater guiding significance. This model can be used not only in the TV series, for the film, network drama, games and other manufacturing can be predicted the future development by this way.

## References

- [1]N. Friedman,D. Geiger,M. Goldszmidt, Bayesian Network Classifiers, Machine Learning . 29(1997) 131–163.
- [2] Y.T. Wang, S.F. Quan, The Current Trends and Thoughts on Science Communication in Science Fiction Film: Based on a Categorical Data Analysis of Top 100 Sci-Fi Film in IMDb,Studies on Science Popularization . 5(2016)54-61.
- [3] S.S. Ma, Bayesian Statistics (Eds.), China Statistics Press, Beijing, 2012.