



Project Name:- Rating Prediction Project

SUBMITTED BY:- VARSHA V. SHINDE



INDEX

SR.NO	TOPIC	PAGE.NO
1	PROBLEM STATEMENT	3
2	EDA STATEMENT & VISUALIZATION	4
3	ADVANTAGES	5
4	EFFECT OF REVIEW & RATING	6
5	CONCLUSION	7

PROBLEM STATEMENT

- Electronic commerce is becoming increasingly popular due to the fact that e-commerce websites allow purchasers to leave reviews on different products. Millions of reviews are being generated everyday by costumers which makes it difficult for product manufacturers to keep track of customer opinions of their products.
- Thus, it is important to classify such large and complex data in order to derive useful information from a large set of data. Classification methods are the way to tackle such problems.
- Classification is the process of categorizing data into groups or classes based on common traits .
- A common concern for organizations is the ability to automate the classification process when big datasets are being used

EDA STATEMENT & VISUALIZATION

- When looking at a data frame, we would think of checking the major feature intuitively, which in this case might be the revenue and the budget. For analyzing the simple variable, we could look at its univariate distribution.
- The histogram is, most of the time, an excellent way to find out the distribution of a single variable. As shown in the figure below, the histogram bins each value and count the occurrence of values within each bin (each bin represent a range of value).
- Thus, it sometimes requires more consideration of the determination of the bin size. By plotting the histogram, we can observe that the distribution of the revenue and the budget are both exponential decaying variable (which means that more movies are with lower budget and revenue).

Scatter matrix

- Different from the scatter plot, scatter matrix plots the scatter plot for every combination of each feature we choose. It gives us a quick check of the correlations between each feature. Specifically, the diagonal scatter matrices represent the histogram of the variable.

ADVANTAGES

- User gets genuine reviews about the product.
- User can post their own review about the product.
- People can easily decide whether the product is good or bad by using this application.
- Since system ranks the feedback based on the weight age of the keywords in database, so the result is appropriate.

EFFECT OF REVIEW & RATING

- THE EFFECT OF REVIEW:-
- Prior research on review volume has produced mixed results. Volume is believed to exert positive effects by increasing review credibility and/or signalling product popularity. Volume has been shown to have a positive effect on box office sales and sales rank of electronic products .
- Moreover, it influences other outcomes such as consumer attention , product evaluation , product popularity and purchase intention . However, finds that volume does not impact the sales of cell phones, and shows that volume has no significant impact on box office performance.
- There is no evidence for a negative effect. In order to understand the effect of volume on purchase we propose the following assumptions:

CONCLUSION

- In this paper, we incorporate sentiment analysis of reviews techniques into the spam review detection. First we have made our own dictionary having sentiment words along with the weight given to the word according to its polarity.
- Then a method is proposed to calculate the sentiment score of the reviews from the natural language text by a shallow dependency parser. A set of discriminative rules are presented through intuitive observation.
- The discriminative rules are combined with the time series method to find the spam and fake reviews. Then the experimented case study and dataset demonstrate the efficiency of our proposed method. In future we would try to improve the method of calculating the sentiment score of the reviews.
- We would also try to update our dictionary containing sentiment word. We would try to add more words in our dictionary and update the weights given to those words to get more accurate calculated score of the reviews.
- Moreover, the effect of valence can be better studied and understood. What is the value of higher ratings versus lower ratings? Studying the interaction between volume and valence can also help with understanding the way that customers interpret consumer reviews. For example, all else being equal, is it better for a product to have a 4.5 star average rating based on 5 reviews or a 4 star average based on 50 reviews?