

## Sentiments' Analysis of Users' Review to Improve 5 Star Rating Method for a Recommendation System

Surabhi Arya, Graduate Student, Oklahoma State University

### ABSTRACT

Recommendation Systems plays a significant role in e commerce industry for making personalized recommendations to users. Users with the help of different rating methods rate items and these ratings are utilized by algorithm to make recommendations to other users. Five-star rating method is a popular platform to record users rating, but it also has some disadvantages which are described in upcoming section of this paper. This paper aims to solve these problems by using sentiment analysis of users' review.

I extracted the data from Amazon.com using web scraping techniques.

### INTRODUCTION

Companies worldwide are using the power of recommendation system to enrich user interaction and their engagement with the websites. Recommendation systems has the potential to change the way companies communicate with users and allow companies to maximize their revenue.

The algorithm behind recommendation system utilizes users rating for different products for ranking and whenever a user makes an attempt to search for any product the system prompts with the top rated products.

There are several rating techniques through which the website can record uses' response in terms of rating. Most popular type of rating methods is five-star method. Following are the examples from some popular e commerce website.



Fig. Shows rating methods used in different websites like Amazon, Airbnb, Zomato (follows order left to right)

But, five-star rating method holds several disadvantages which doesn't allow recommendation system to capture enough information. These problems are explained as below:

1. Bimodal Distribution: Based on the study of Amazon data, it has been observed that these ratings shows bimodal distribution that means rating tends to cluster round extreme numbers which in our case is 1 and 5. Thus averaging all the ratings from users for a particular product and rank that product is not an accurate reflection of product quality.
2. Missing important details: Star rating system is unable to explain important details or properties about product. For instance, Product A can be beneficial for user X in a particular set of environment and it may or may not be equally beneficial in other set of environment.

This paper intends to solve above problems by doing sentiment analysis of users' review using SAS EM Text Miner software.

## METHODOLOGY

### 1. Data Description

The idea behind doing sentiment analysis for users' review is to extract their feeling about product to provide a better rating insights. In order to capture sentiments about a product, I have used data from [https://www.amazon.com/All-New-Amazon-Echo-Dot-Add-Alexa-To-Any-Room/product-reviews/B01DFKC2SO/ref=cm\\_cr\\_getr\\_d\\_paging\\_btm\\_next\\_10?ie=UTF8&reviewerType=all\\_reviews&filterByStar=positive&pageNumber=10](https://www.amazon.com/All-New-Amazon-Echo-Dot-Add-Alexa-To-Any-Room/product-reviews/B01DFKC2SO/ref=cm_cr_getr_d_paging_btm_next_10?ie=UTF8&reviewerType=all_reviews&filterByStar=positive&pageNumber=10), using web scraping techniques.

Data was extracted in two segments namely top positive review and top critical review having 1k observations in each segment.

Following table shows metadata for the data used in analysis

| Column Name | Type | Description  | Role |
|-------------|------|--|------|
| ID          | Num  | It is a unique identification number for each user | ID   |
| Reviews     | Char | Product reviews of users' obtained from websites   | Text |

### 2. Analysis

The analysis was done using Text Mining using SAS EM Text Miner Software. Unstructured data was converted into structured data identifying frequently occurring words and important topics on which comments were being made.

1. Text Parsing Node - This node creates term by frequency matrix of unstructured text data. The terms identified in this matrix are further passed into text filter node.
2. Text Filter Node – This node is added to text parsing node and is used to eliminate terms which appears least number of times in the document by changing the term filter setting in property panel.
3. Concept Link – This helps in understanding the relationship between words (terms). This link chart is obtained in filter viewer option in property panel.
4. Text Topic - Topics are terms that describes the main idea of your data, text topic node enables you to explore the document by relating terms and documents based on discovered and user defined topics.

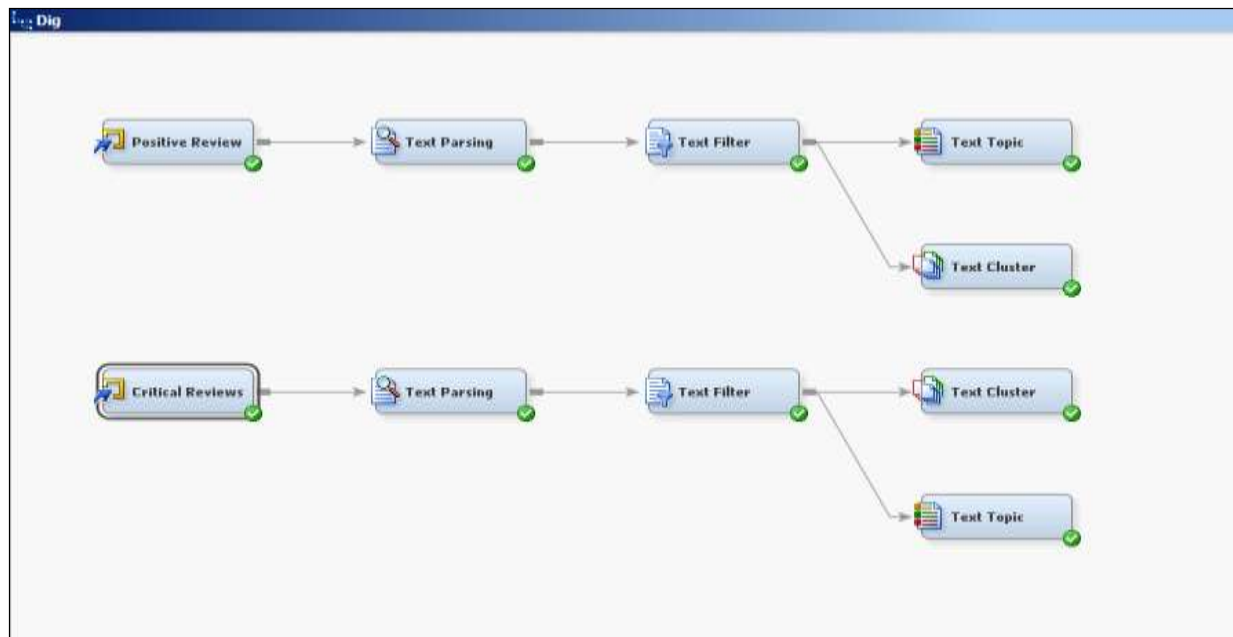
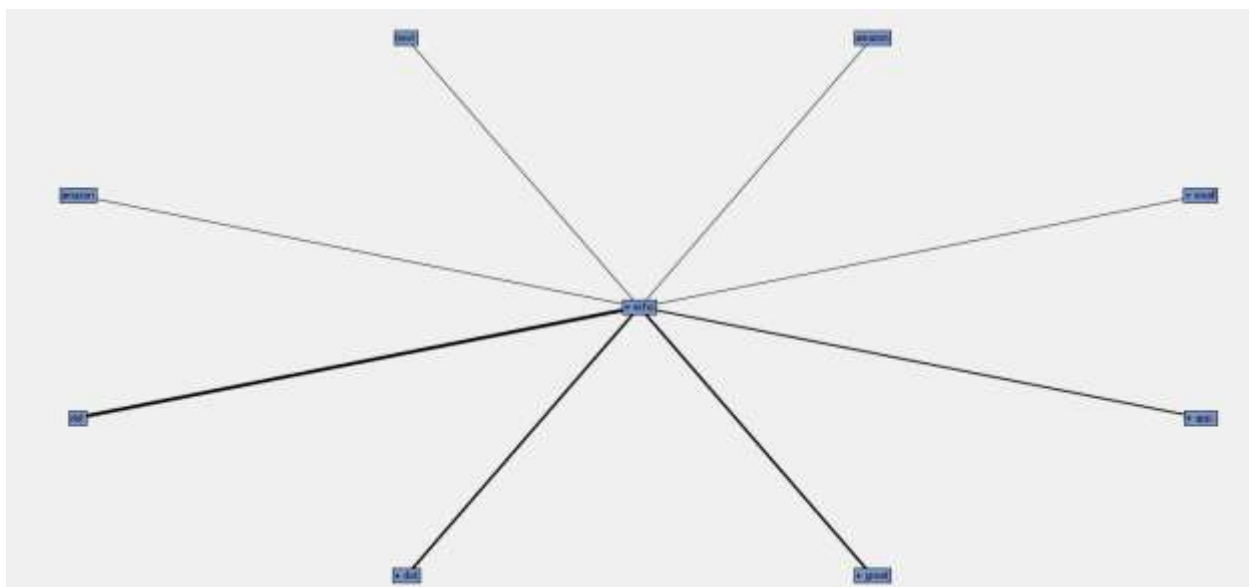


Fig. Shows SAS EM Diagram for process

## Echo Dot (2<sup>nd</sup> Generation) – Smart Speaker with Alexa – Black Reviews

### Positive Reviews



Line connecting echo with different terms represents the degree of association, which implies echo is highly correlated to dot, +dot, +great and +app terms. Further to know, how these terms are further associated with other terms, link graph is extended to more terms.



The concept link shown above represents high correlation between +echo and other terms like +issue, +dot. Users are facing problems related to software like unrestricted Bluetooth pairing, all the contact of the phone are automatically stored to Alexa cloud etc.

### 3. Discussion

Among all the topics identified, only significant topics are picked for the analysis.

Topics building for positive reviews

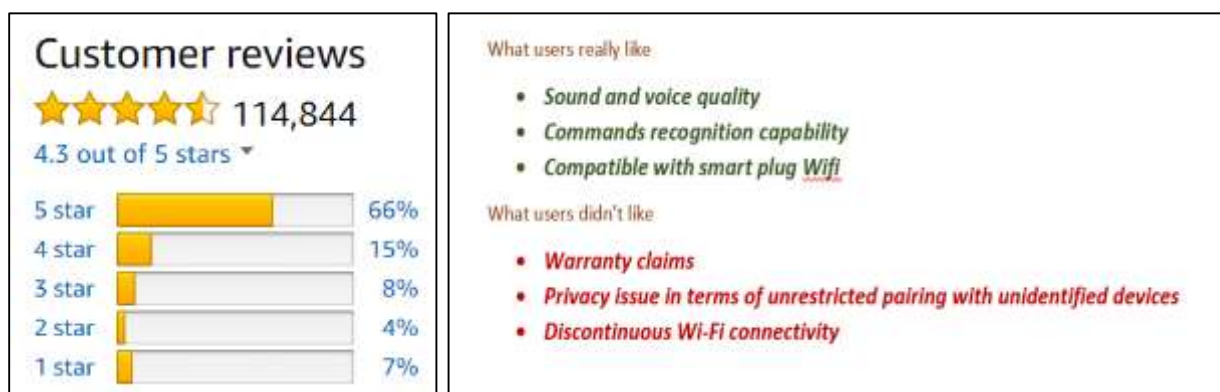
| Topics                                      | Explanation   |
|---|---|
| +day,+time,+set,+remind,+love               | People really loved features like timer, reminder, weather details  |
| +sound,+speaker,+song,+buy,+order           | People enjoyed listening music with its clear and loud sound quality  |
| Siri,+understand,tv,+price                  | Users found it comparable to Siri, impressed by Alexa understanding capability that it does not pick up voice from TV |
| Wifi,+smart plug,quality,Christmas, turn on | Compatibility with smart plug Wifi is a likable feature and people prefer to purchase during Christmas sale           |

Topics building for negative reviews

| Topics                                      | Explanation   |
|---|---|
| +warranty,+technical,+item,+purchase        | People had complaints about warranty of the product when they requested to return the device                      |
| +dot,+connect,security,Bluetooth, neighbour | Users are unhappy about unrestricted pairing of any device in the vicinity of the device, hindering their privacy |
| Support,tech,+contact,+require,+replace     | Users had complaint about Amazon customer support involving discontinuous Wi-Fi connectivity                      |

## RESULTS

By doing text analysis, I was able to identify highly frequent sentiments from users' review which were not captured by 5-star rating before. One can leverage these sentiments along with ratings to provide better insights about the product. Following figure demonstrates the improved rating method.



This method can be further used to make better recommendations. When a user attempts to search for a product and service, the recommendation prompts with the top rated products. If we provide user, products sentiments along with the users rating, it will be easier for user to decide what is the best option according the requirement and thus solves the problem of missing important information and bimodal distribution

## FURTHER SCOPE

The text analytics explained in this paper can be used for different websites like Zomato, IMDB where customers' credibility towards the website is dependent upon the reviews and rating provided by the other users.

## CONCLUSION

This paper provides a new insight on how text mining leading to sentiments analysis can improve commonly used five-star rating method. Along with that we also understood basics text parsing, text topic and concept link. I took the instance of Echo Dot (2<sup>nd</sup> Generation) – Smart Speaker with Alexa – Black and analyzed users' review to extract significant sentiments for this product. As a result, I was able to capture sentiments about what users really like and what not which is further integrated with five-star rating system.

## REFERENCES

Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS® by Goutam Chakraborty, Murali Pagolu, Satish Garla. 3) Sentiment Analysis and Opinion Mining by Bing Liu

Sentiment Analysis on YouTube Movie Trailer comments to determine the impact on Box-Office Earning Rishanki Jain, Oklahoma State University