International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 2, January-February 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Study and Analysis of Online Comment Data Mining and Kano Model Research

Nilza Angmo[1], Er. Vandana[2]

[1]M Tech Scholar, [2]Assistant Professor,
[1,2]Department of Computer Science Engineering, SVIET, Banur, Punjab, India

ABSTRACT

The foundation of this study is based on the Kano model. This model was initially used by making a questionnaire survey. But now a days due to vast amount of data and opinions available via the internet on the World Wide Web the model can be changed. In this study data from an e-commerce site has been collected and has undergone preprocessing, sentiment analysis and then Kano evaluation to understand how to satisfy a customer. The use of the said model overcomes the disadvantages of the sentiment analysis. In this study a combination of sentiment analysis and Kano model has been done on online comment data giving us picture of how to discover the demands of a customer and also how to satisfy him.

KEYWORDS: *Data Mining, Kano Model, Big Data, Social Media*

How to cite this paper: Nilza Angmo | Er. Vandana "Study and Analysis of Online Comment Data Mining and Kano Model Research" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-2, February 2021, pp.636-641, URL: www.ijtsrd.com/papers/ijtsrd38528.pdf

IJTSRD38528

INTRODUCTION:

There has been a great transformation in the way we associate with the world due to social media. There is a trend wherein various social media platforms like Facebook and Twitter and used by individuals to find and share data and interact with people. This has resulted in the generation of a lot of opinions online. These opinions further provide help in anticipating the future as well as investigate present. The collection of data for the investigations on social media is done via various platforms like Facebook, Twitter, and LinkedIn etc. The impact of social media is not just on the way that people interact with one another but has also resulted in reimagining the way in which the businesses are planned by various organizations. Social media has found its way of making businesses more successful. This is due to the sharing of reviews and assisting buyers to make a better choice. It also provides the sellers with an idea of where a certain product or service lags and thus allows them to improve upon it in order to gain a customer's trust and loyalty.

LITERATURE REVIEW:

WuBi in the research paper [1] describes the increasing number of online reviews and importance of such online posted reviews to understand the customer satisfaction. The paper proposed a model for customer satisfaction from online reviews. In the research paper [2] *Jin* presents the comprehensive review information mining from big consumer opinion data to assist product design. Reviews on information utilization of big consumer opinion data for product design are explored in terms of how to extract critical customer needs from big consumer opinion data, how to connect the voice of the customers with product design, how to make effective comparisons and reasonable ranking on similar products, how to identify ever-evolving customer concerns efficiently, and so on. Furthermore, significant and practical aspects of research trends are highlighted for future studies. In paper [3] *Song and Chen* describe the key points of the text analysis in terms of opinion mining and the sentiment analysis. In this paper, we come up with the model to combine the data mining technology with the Kano model, first we discover the feature theme of the product by establishing the comment mining model, and analysis the sentiment of the comment through machine learning to acquire the parameters of Kano model such as the initial importance. In research paper [7] by Rotar identification of underlying home appliance factors were identified and the results were used to construct the Kano model. This study helped optimize business decision making with scientific research. In paper [4] the sentiment analysis of microblogs and its importance in academic and Industrial fields is described. The paper simply presents an opinion mining system for Miner a Chinese microblogs. The paper [5] states that twitter data contains many sentiments which can be analysed using Hadoop. Twitter's API is used to derive data from twitter. Then the data undergoes some processing. First removal of stop words is done. Then the tokens are changed into a structured form as they are mostly in unstructured form. Then emoticons are also translated for

higher accuracy. Then Map-Reduce is used to find sentiment of each word and the sum gives the overall sentiment of the tweet. In paper [6] first map-reduce, a sentence is detected and stopwords, hashtags etc. are removed. Then we search for words which represent features and are then clustered. Then Open NLP is used for POS Tagging. Phrase removal is done before stop word removal. In second map-reduce, a sentiwordnet dictionary is used, scores are given to words, which is averaged then to get overall value. *Priya. V* has examined the sentiment of youngsters regarding the floods in Chennai in 2016. They used flume to get data from twitter and applying Naïve Bayes algorithm. They developed a dictionary to compare the tweets with and get a sentiment score [8]. *M. Edison* [9] showed various methods and concepts of sentiment analysis on big data. It has two approaches- Lexicon based and Machine Learning based. The machine learning approach is more popular and uses various supervised and unsupervised learning algorithms. In [10] use of Weka, on open source tool for data mining, to perform sentiment analysis for movie reviews. Data was taken from twitter and other online review platforms like IMDB. Then data needs to be pre-processed. Then naive bayes classifier is used. Then the accuracy needs to be checked.

PROPOSED WORK:

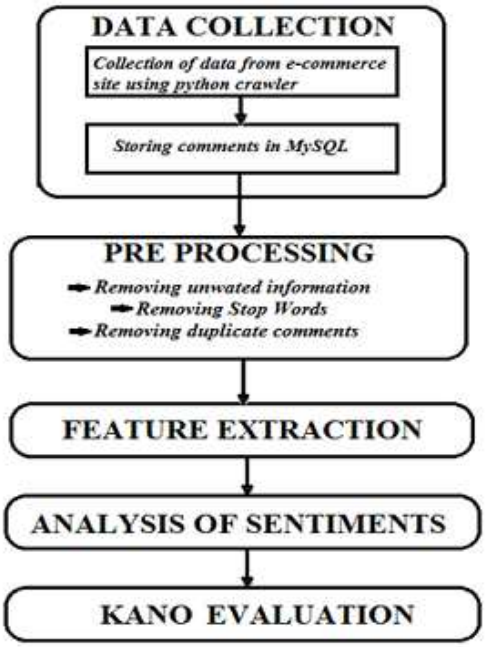


Figure 3.1: Flow Chart of the Proposed Methodology.

RESULTS:

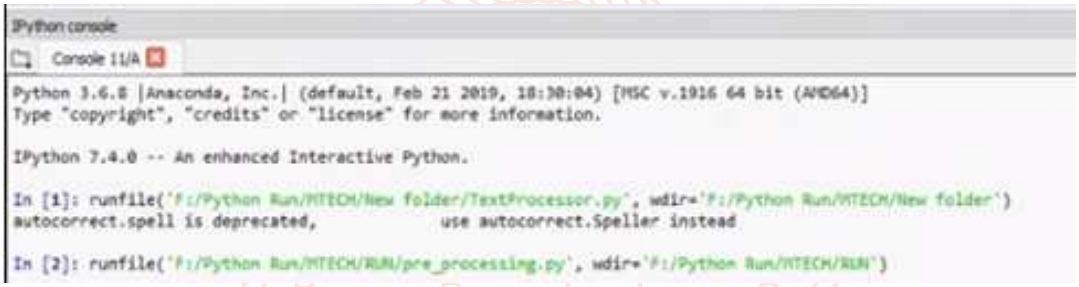


Figure 1: Pre-Processing the data.

Stemming words: In this process retrieval of information the words are reduced to their stem or base or root.

Stopping Words: Stop words are words which are filtered out prior to or after processing of natural language data.

Comments Pre-Processing: The comments are preprocessed. After this the text processing is done on all the comments that have been pre-processed.

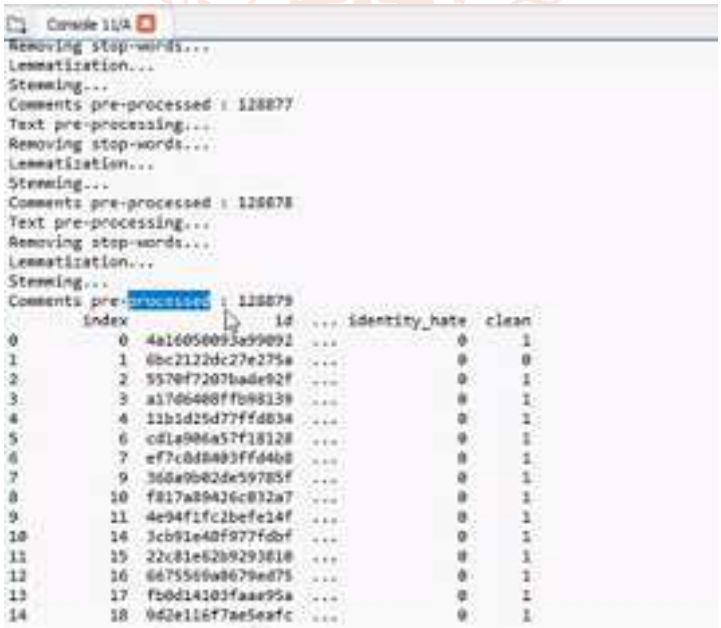


Figure 2: Removal of stop words, Lemmatization, Stemming Followed by pre processing of Comments.

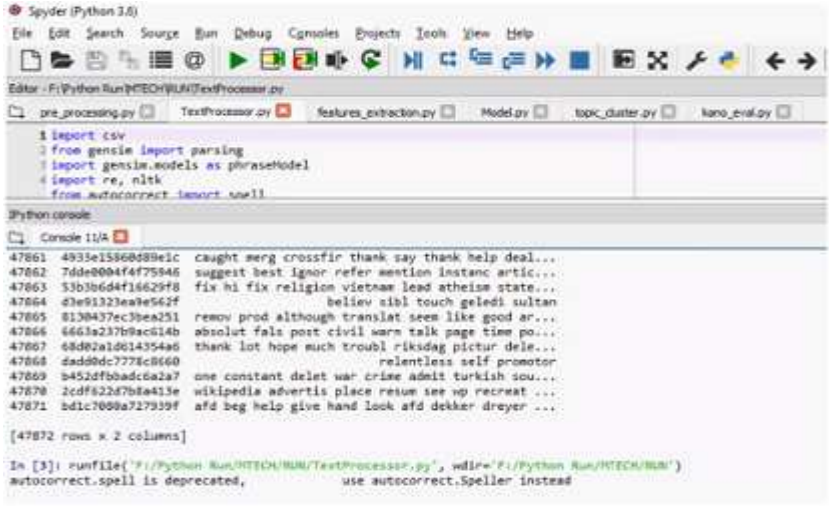


Figure 3: Text-Processing.

The preprocessing stage is followed by feature extraction. A filter is applied to get the desired data. This is required later in order to create training data sets for the model.

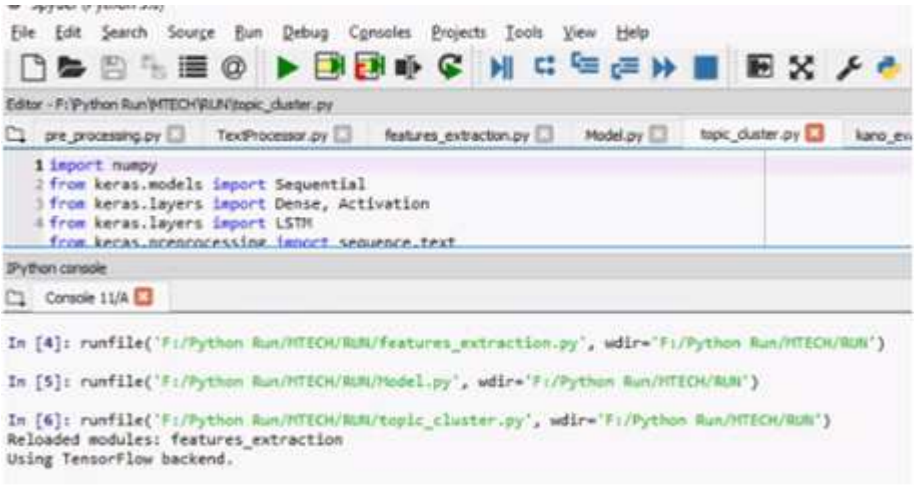


Figure 4: Feature Extraction and Clustering.

Topic Cluster:

Creating text clusters in order to perform unsupervised learning a clustering algorithm is used. The code is run five times. It simply determines the clusters. The data is then fed to the model created in order to train it and then perform sentiment analysis on a set of records. The Accuracy is computed by checking the correct predictions by total number of predictions. The figures that follow show the graphs.

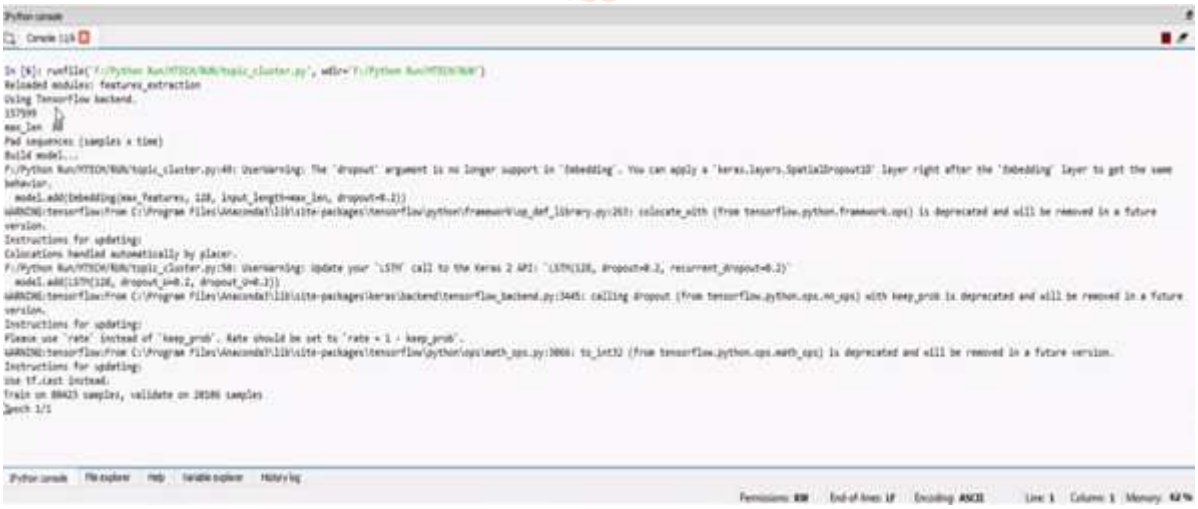
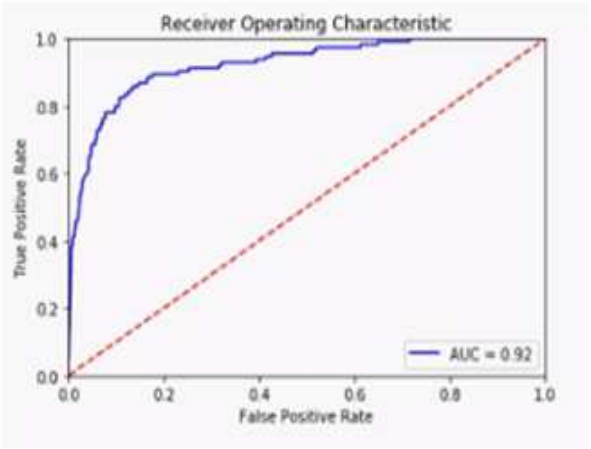
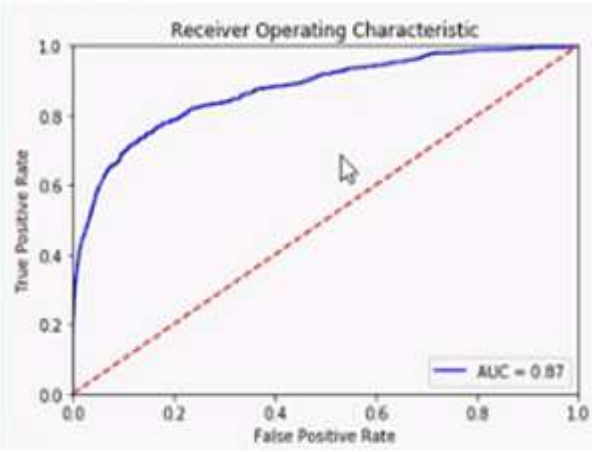


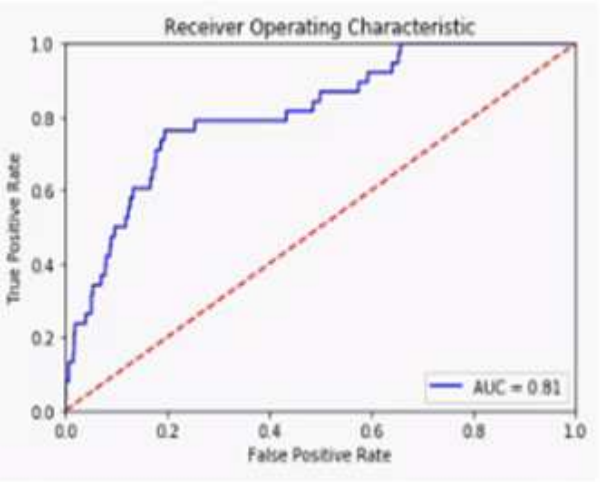
Figure 5: Topic Cluster.



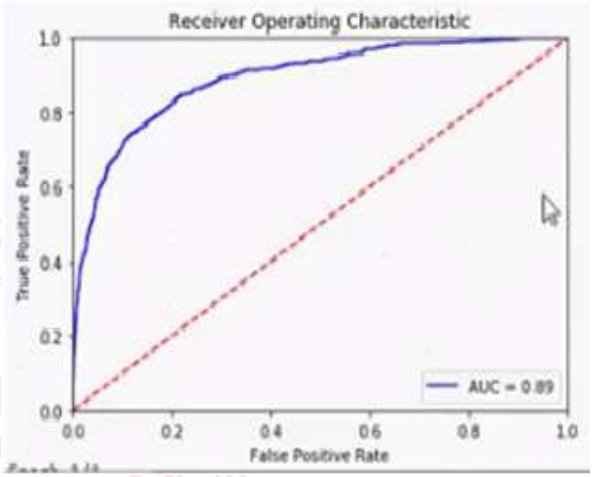
Graph 1:Run1- Prediction model with accuracy of 0.92.



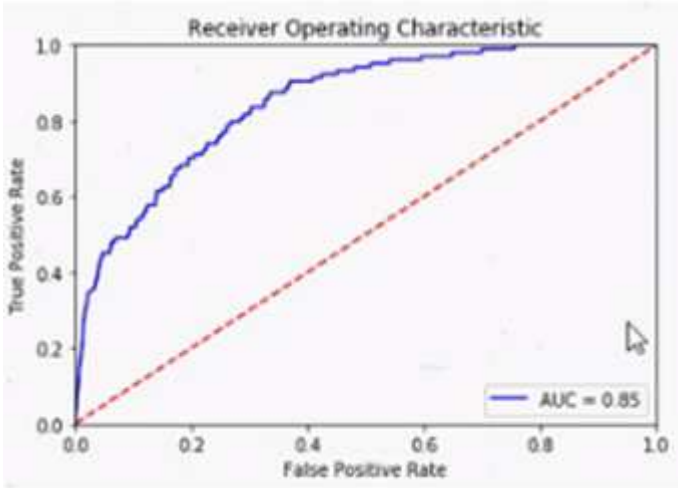
Graph 2: Run2-Prediction model with accuracy of 0.87.



Graph 3:Run3-Prediction model with accuracy of 0.81.



Graph 4:Run4-Prediction model with accuracy of 0.89.



Graph 5:Run5-Prediction model with accuracy of 0.85.

Kano Evaluation: The Kano evaluation response graphs are shown in the figures that follow:

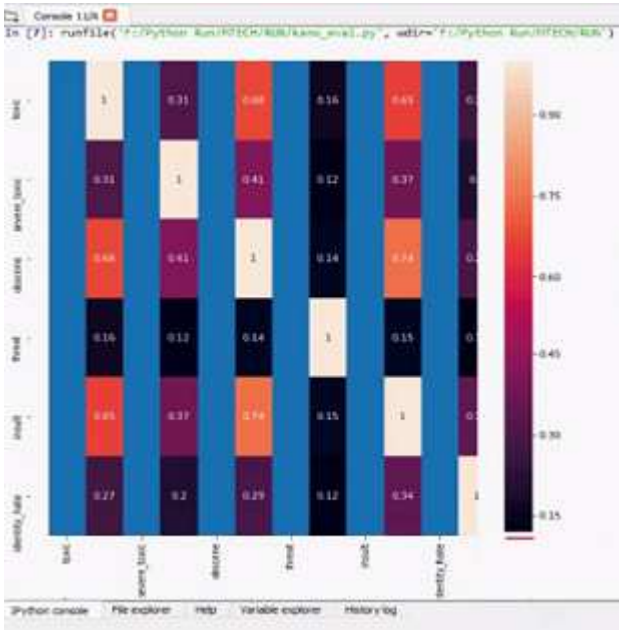


Figure 4.13: Color coded response matrix.

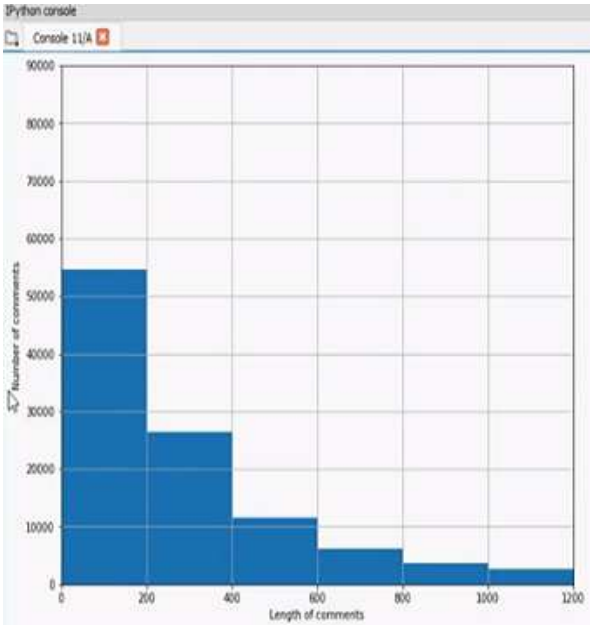


Figure 4.13: Processed vs. Length

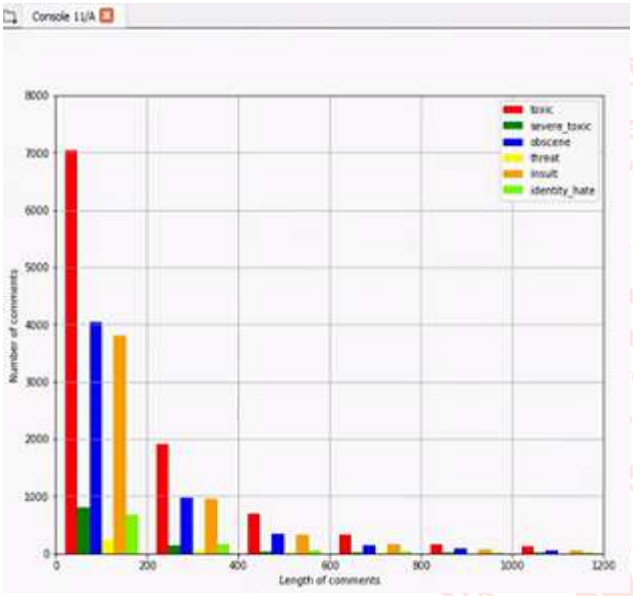


Figure 4.14: Polarity of each comment.

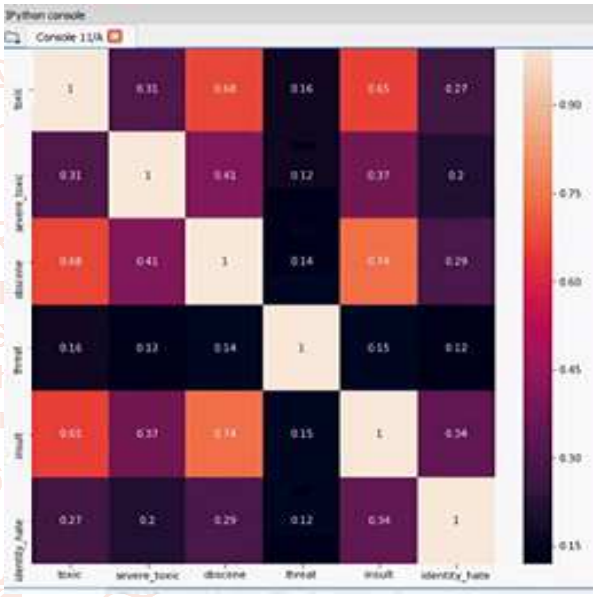


Figure 15: Color Matrix.

	Must-have	Linear	Exciter	Reverse	Questionable	Inverse
The solution allows you to still communicate with others	14	10	14	1	0	1
The solution provides you smarter information about your phone's battery performan	7	3	15	1	0	20
The solution provides options to charge your phone whenever you are	4	10	21	1	0	8
The solution is convenient and helps you maintain your productivity	11	13	15	0	0	9
The solution doesn't upset the design aesthetic of the actual phone	10	4	11	0	1	22
The solution makes you feel safe	8	5	11	1	0	25

Figure Results.

CONCLUSION:

This dissertation takes the review of a product online to analyze the sentiment of the posted comment using machine learning approach and by using the data mining and Kano model evaluate the same. This proposed model can be integrated with QFD to overcome the drawbacks of the old methods used for the same. Previously, the Kano model get client needs data from the poll/questionnaires, which was small and subjective. The proposed system is based on reviews of people online, through the foundation of a review mining model to distinguish the qualities of the item traits, and the item highlights and remarks are mapped to the feature space to examine the statement of the customer's enthusiastic worth. This model will thus prove to be very useful.

REFRENCES

[1] Jian-Wu Bi, Yang Liu, Zhi-Ping Fan & Erik Cambria, "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model", 2019

[2] Jian Jin, Ying Liu, Ping Ji, C.K. Kwong, "Review on Recent Advances in Information Mining From Big Consumer Opinion Data for Product Design", 2018, Journal of Computing and Information Science in Engineering, Volume 19, Issue 1.

[3] Huaming Song, Chao Chen, "Research on Kano Model Based on Online Comment Data Mining", 2018, IEEE.

[4] Xinjie Zhou, Xiaojun Wan, and Jianguo Xiao, "CMiner: Opinion Extraction and Summarization for Chinese Microblogs", 2016, IEEE.

[5] Divya Sehgal and Dr. Ambuj Kumar Agarwal, "**Sentiment Analysis of Big Data Applications using Twitter Data with the Help of HADOOP Framework**", IEEE, 5th International Conference on System Modelling & Advancement in Research Trends, 25th -27th November, 2016, pp. 251-255.

[6] Jalpa Mehta, Jayesh Patil, Rutesh Patil, Mansi Somani and Sheel Varma, "**Sentiment Analysis on Product Reviews using Hadoop**", International Journal of Computer Applications Volume 142 – No.11, May 2016, pp. 38-41

[7] Laura Južnik Rotar, Mitja Kozar, "The Use of the Kano Model to Enhance Customer Satisfaction", 2017 DE GRUYTER Open. Vol 50.

[8] Priya. V, S Divya Vandana, **"Chennai Rains Sentiment-An Analysis Of Opinion About Youngsters Reflected In Tweets Using Hadoop"**, International Journal of Pharmacy & Technology, Sep-2016, Vol. 8, Issue No.3, pp. 16172-16180.

[9] M. Edison, A. Aloysius, **"Concepts and Methods of Sentiment Analysis on Big Data"**, International Journal of Innovative Research in Science Engineering and Technology, Vol. 5, Issue 9, September 2016, pp. 16288-16296.

[10] Rajni Singh and Rajdeep Kaur, **"Sentiment Analysis on Social Media and Online Review"**, International Journal of Computer Applications, July 2015, Volume 121, Issue 20, pp. 44-48.