International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Aspect Based Sentiment Analysis On Amazon Earphones Sales

Chavva Pooja[1], B.Charan Babu[1], G.Satish Kumar[1], B.Venkateswarlu[1], G.Manthru Naik[2]

[1]UG Students, [2]Associate Professor, [1,2]Department of Computer Science Engineering, Kallam Haranadha Reddy Institute Of Technology, Guntur.

ABSTRACT:

The main objective of this project is to present an aspect and description system based sentimental analysis for Amazon sales data on different manufactures including boat, Samsung etc… The computational handling of views, feelings, and subjectivity in text is known as sentiment analysis (SA). Aspect-based Sentiment Analysis (ABSA) is a type of sentiment analysis that tries to identify the most significant characteristics of an entity from the text and forecast their polarity. An examination of ABSA's most current state reveals a significant increase in identifying both aspect and emotion. The suggested algorithms and models are used to classify current techniques. The dataset, domain, and reported performance are all given for each mentioned study, as well as the aspect extraction method and sentiment prediction approach. The main objective of this project is to provide a quick overview of ABSA methods. The task's primary contributions are enhanced categorizations of a large number of recent Amazon earphone sales based on manufacturers such as boat, Samsung, and others.

Keywords: ABSA, characteristics, performance, sentiment analysis.

1.Introduction:

Aspect sentiment analysis is essential because it may assist businesses in automatically sorting and analyzing client data, automating procedures such as customer care activities, and gaining useful insights on the go. Customers are more outspoken than ever before. They like interacting with businesses and providing feedback, both positive and negative. Customers leave a lot of information every time they contact with a company, whether it is through a mention or a remark, letting businesses know whether they are doing well or wrong. However, shifting through all this data by hand may be challenging.Aspect-based sentiment analysis, on the other hand, does the heavy lifting for you. The computational study of people's views, attitude, emotion, or evaluation relating to subjects, items, goods, services, organizations, persons, and events, or their qualities, is known as sentiment analysis. Researchers in the disciplines of data mining and natural language processing have been paying close attention to sentiment analysis for online consumer evaluations in recent years. Traditional document-level sentiment classification attempts to determine if a particular text's general sentiment polarity is positive, negative, or neutral.To extract sentiment for each aspect from the text corpus, the ABSA model requires aspect categories and their associated aspect words. For a

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

given implementation, a domain-specific model can be created; however, general language models can also be utilized. Each statement must include labeled data comprising aspect words and aspect categories, as well as a sentiment score. It can, however, be solved without labeled data or a list of aspect words using the unsupervised technique. This is the reason that a new field of opinion mining or Sentiment Analysis (SA) has emerged. Liu (2012) define SA as "the field of study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes." Document-level and sentence-level sentiments cannot provide sufficient information that is important for decision making. We can obtain such information with aspect level SA. If a reviewer gives a feedback on a particular product, he usually comments on some aspects of that product.

2.Literature Review:

[1]Sentiment Analysis:An Enhancement of Ontological-Based Using Hybrid Machine Learning Techniques.With the fast development of World Wide Web 2.0 has resulted in huge number of reviews where the consumers share their opinion about a variety of products in the websites, forum and social media such as Twitter and Instagram. For the organizations, they have to analyze customer's behavior to find new market trends and insights. Sentiment analysis concept used to extract the positive, negative or neutral sentiment of the features from the unstructured data of product reviews. In this paper, we explore the techniques and tools used to enhance the ontology- based approach. Combination of ontology-based on Formal Concept Analysis (FCA) which a process of obtaining a formal ontology or a concept hierarchy from a group of objects with their properties and K-Nearest Neighbor (KNN) to classify the reviews. **Summary:** Abu Latiffi, M. I. & Yaakub describes in this paper, how they are able to view the strength and weakness of the product in more detail where the feature selection process will more be systematic and will result in the highest feature set.

[2] An Aspect Detection Model Based on Topic Modeling Using the Structure of Review Sentences:

Probabilistic topic models are statistical methods whose aim is to discover the latent structure in a large collection of documents. The intuition behind topic models is that, by generating documents by latent topics, the word distribution for each topic can be modeled and the prior distribution over the topic learned. In this paper we propose to apply this concept by modeling the topics of sentences for the aspect detection problem in review documents in order to improve sentiment analysis systems. Aspect detection in sentiment analysis helps customers effectively navigate into detailed information about their features of interest. **Summary**: Bagheri, A., Saraee, M. & De Jong, F, observes that the model is indeed able to perform the task significantly better when compared with standard topic models.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

[3] Building a Sentiment Summarizer for Local Service Reviews. Proceedings of WWW Workshop on NLP in the Information Explosion: Online user reviews are increasingly becoming the de-facto standard for measuring the quality of electronics, restaurants, merchants, etc. The sheer volume of online reviews makes it difficult for a human to process and extract all meaningful information in order to make an educated purchase. As a result, there has been a trend toward systems that can automatically summarize opinions from a set of re- views and display them in an easy to process manner. In this paper, we present a system that summarizes the sentiment of reviews for a local service such as a restaurant or hotel. In particular we focus on aspect-based summarization models, where a summary is built by extracting relevant aspects of a service, such as service or value, aggregating the sentiment per aspect, and selecting aspect-relevant text. **Summary:** These models are unique in that they leverage user-supplied labels and domain-specific features of service evaluations to improve quality.

3.Proposed System:

3.1. We propose this application that can be considered a useful system since it helps to reduce the limitations obtained from traditional and other existing methods. The objective of this study is to develop fast and reliable method which detects the sentiment accurately. To design this system is we used a powerful algorithm in a Python based environment.

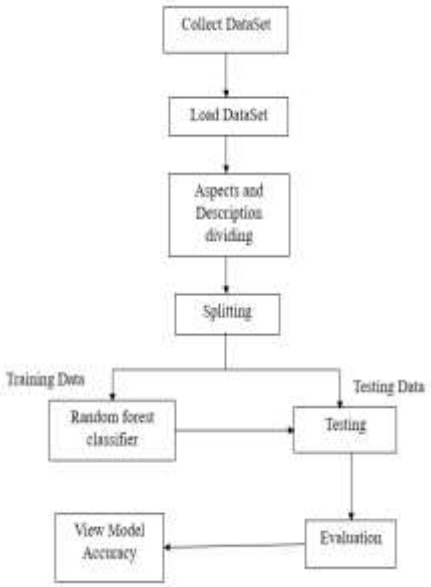

Fig 1:Block Diagram of Proposed system

3.2.Methodology:

TextBlob:

Natural Language Processing (NLP) is an area of growing attention due to increasing number of applications like Chabot's, machine translation etc. In some ways, the entire revolution of intelligent machines in based on the ability to understand and interact with humans. We have been exploring NLP for some time now. Our journey started with NLTK library in Python, which was the recommended library to get started at that time. Later, we got introduced to TextBlob, which is built on the shoulders of NLTK and Pattern. A big advantage of this is, it is easy to learn and offers a lot of features like sentiment analysis, pos-tagging, noun phrase extraction, etc. It has now become my go-to library for performing NLP tasks. On a side note, there is spacy, which is widely recognized as one of the powerful and advanced library used to implement NLP tasks.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

But having encountered both spacy and TextBlob, I would still suggest TextBlob to a beginner due to its simple interface.Installation of TextBlob in your system in a simple task, all you need to do is open anaconda prompt (or terminal if using Mac OS or Ubuntu) and enter the following commands:

pip install -U textblob

This will install TextBlob. For the uninitiated – practical work in Natural Language Processing typically uses large bodies of linguistic data, or corpora. To download the necessary corpora, you can run the following command.

python -m textblob.download_corpora

Tokenization refers to dividing text or a sentence into a sequence of tokens, which roughly correspond to "words". This is one of the basic tasks of NLP. To do this using TextBlob, follow the two steps:

1. Create a textblob object and pass a string with it.
2. Call functions of textblob in order to do a specific task.

Part-of-speech tagging or grammatical tagging is a method to mark words present in a text on the basis of its definition and context. In simple words, it tells whether a word is a noun, or an adjective, or a verb, etc.

The *sentiment* function of textblob returns two properties, polarity, and subjectivity.

- Polarity is float which lies in the range of [-1,1] where 1 means positive statement and -1 means a negative statement. Subjective sentences generally refer to personal opinion, emotion or judgment whereas objective refers to factual information. Subjectivity is also a float which lies in the range of [0,1].

Random Forest :

A random forest is a machine learning technique that's used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.

A random forest algorithm consists of many decision trees. The 'forest' generated by the random forest algorithm is trained through bagging or bootstrap aggregating. Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms.

Features of a Random Forest Algorithm:

- It's more accurate than the decision tree algorithm.
- It provides an effective way of handling missing data.
- It can produce a reasonable prediction without hyper-parameter tuning.
- It solves the issue of over fitting in decision trees.
- In every random forest tree, a subset of features is selected randomly at the node's splitting point.

For earphones the root nodes features that could influence the customer's choice (price, sound quality). The random forest will split the nodes by selecting features randomly. The final prediction will be selected based on the outcome of the four trees.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

4.Results:

Count plot, which shows the count of ratings with respect to the products.

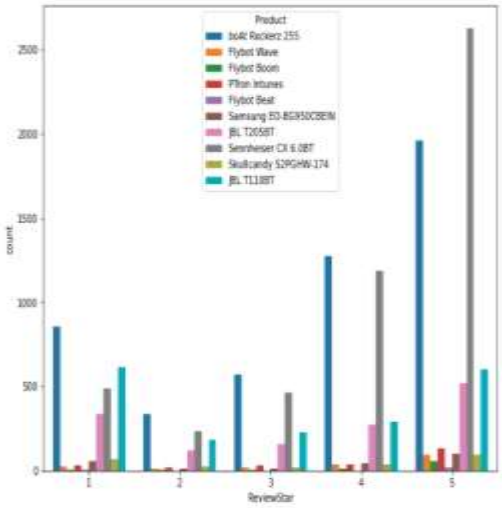

Fig:Count Plot

Info, gives us an overview of dataset .

```
df.info
```

```
<bound method DataFrame.info of                                        ReviewTitle ...        Product
0              Honest review of an edm music lover\n ...  boAt Rockerz 255
1            Unreliable earphones with high cost\n ...  boAt Rockerz 255
2                    Really good and durable.\n ...  boAt Rockerz 255
3                stopped working in just 14 days\n ...  boAt Rockerz 255
4     Just Awesome Wireless Headphone under 1000...😍\n ...  boAt Rockerz 255
...                                                 ...  ...          ...
14332                                    Good\n ...  JBL T110BT
14333                            Amazing Product\n ...  JBL T110BT
14334                                 Not bad\n ...  JBL T110BT
14335                              a good product\n ...  JBL T110BT
14336        Average headphones , n overrated name\n ...  JBL T110BT

[14337 rows x 4 columns]>
```

Fig: Dataset overview

Resulted sentiment from the outcomes

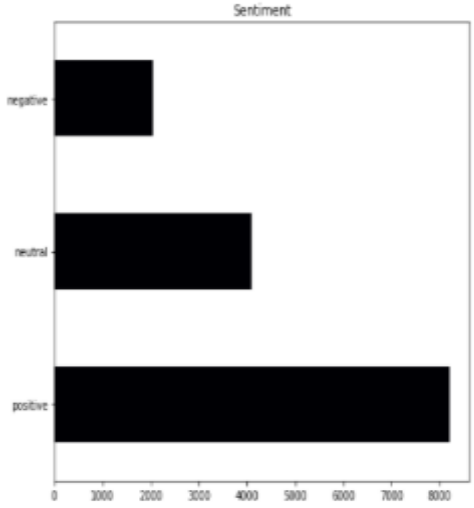

Fig: Sentiment Analysis

5.Conclusion:

We have successfully developed a system which gives the sentiment with respect to the sentiment. This is created in a user-friendly environment with Python programming on Jupyter notebook. The ability to identify numerous reviews could be added to this in the future. We intend to investigate prediction approach with the revised data set and employ the most accurate and relevant machine learning algorithms for detection.

6.References:

[1]Abu Latiffi, M. I. & Yaakub, M. R. 2018. Sentiment Analysis: An Enhancement of Ontological-Based Using Hybrid Machine Learning Techniques. Asia-Pacific Journal of Information Technology and Multimedia. 7(2):61-69.

[2]Bagheri, A., Saraee, M. & De Jong, F. 2014. ADM-LDA: An Aspect Detection Model Based on Topic Modelling Using the Structure of Review Sentences. Journal of Information Science. 40(5): 621-636.

[3]Blair-Goldensohn, S., Hannan, K., Mcdonald, R., Neylon, T., Reis, G. A. & Reynar, J. 2008. Building a Sentiment Summarizer for Local Service Reviews. Proceedings of WWW Workshop on NLP in the Information Explosion Era. pp. 1-10.

[4]Brody, S. & Elhadad, N.2010. An Unsupervised Aspect-Sentiment Model for Online Reviews. Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistic. 804-812.

[5]Shruti Kaushik, Mehul P. Barot, "Sarcasm Detection In Sentiment Analysis", IJARIIE, 2016, ISSN-2395-4396.

[6]Cambria, E., Schuller, B., Xia, Y. & Havasi, C. 2013.New Avenues in Opinion Mining and Sentiment Analysis, "in IEEE Intelligent Systems, 28(2):15-21. Doi: 10.1109/MIS.2013.30.

[7]Kiritchenko, S., Zhu, X., Cherry, C. & Mohammad, S. 2014. Nrc-Canada-2014: Detecting Aspects and Sentiment in Customer Reviews. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). pp. 437-442.

[8]Lal, M. & Asnani, K. 2014. Aspect Extraction & Segmentation in Opinion Mining. International Journal Of Engineering And Computer Science 3(5): 5873-5878. Landauer, T. K., Mcnamara, D. S., Dennis, S. & Kintsch, W. 2013. Handbook of Latent Semantic Analysis, New York.

[9]Taylor & Francis. Lek, H. H. & Poo, D. C. 2013. Aspect-Based Twitter Sentiment Classification. Tools with Artificial Intelligence (ICTAI), 2013 IEEE 25th International Conference on, pp. 366-373.

[10]Li, F., Han, C., Huang, M., Zhu, X., Xia, Y.-J., Zhang, S. & Yu, H. 2010. Structure-Aware Review Mining and Summarization. Proceedings of the 23rd International Conference on Computational Linguistics, pp. 653-661.

[11]Liu, B. 2012. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. 5(1): 1-167.

[12] Liu, B., Hu, M. & Cheng, J. 2005. Opinion observer: analyzing and comparing opinions on the Web. In Proceedings of the 14th international conference on World Wide Web (WWW '05).ACM, New York, NY, USA, pp.342351.
DOI:https://doi.org/10.1145/1060745.1060797.

[13]Liu, K., Xu, L .& Zhao, J.2012.Opinion Target Extraction Using Word-Based Translation Model. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 1346-1356.

[14] Liu, P.,Qian, X. & Meng, H. 2014.Seeing Stars from Reviews by a Semantic-based Approach with Map Reduce Implementation. ASEBigData/ Social Informatics/ PASSAT/ BioMedCom 2014 Conference, Harvard University, December 14-16, 2014, pp. 1-5.

[15] Liu, P., Joty, S. & Meng, H. 2015. Fine-Grained Opinion Mining with Recurrent Neural Networks and Word Embeddings. Conference on Empirical Methods in Natural Language Processing (EMNLP 2015), pp. 1433-1443.

[16]Lizhen, L., Wei, S., Hanshi, W., Chuchu, L. & Jingli, L.2014. A Novel Feature-Based Method for Sentiment Analysis of Chinese Product Reviews. China Communications, 11(3): 154-164.