Sentiment Analysis of E-Commerce Website Based on Review Mining

Rajesh B. Khavane, Pushkraj S. Khairnar, Vaishali C. Thakare, Purva D. Datir

UG Students, Department of Computer Technology

K. K. Wagh Polytechnic, Nashik, Maharashtra, India

Abstract*: Sentiment analysis has become an increasingly popular technique for analyzing customer feedback and product reviews on e-commerce websites. In this study, we explore the application of sentiment analysis to an e-commerce website in order to identify customer sentiment and opinion towards various products and aspects of the shopping experience. Using a machine learning approach and natural language processing techniques, we analyze a dataset of customer reviews to classify them as positive, negative, or neutral. We also examine the most commonly occurring words and phrases within each sentiment category to gain further insights into customer sentiment. Our results show that sentiment analysis can provide valuable insights into customer sentiment and opinion towards various products and aspects of the shopping experience. We demonstrate the potential for e-commerce websites to use sentiment analysis to make data-driven decisions about product development, marketing strategies, and customer service initiatives, ultimately leading to increased customer satisfaction and success in the highly competitive online marketplace.*

Keywords: Sentiment analysis, Mining, Natural Language Processing

I. INTRODUCTION

Sentiment analysis is the process of analyzing text data to determine the emotional tone and subjective opinions expressed within it. In the context of an e-commerce website, sentiment analysis can be used to mine product reviews and customer feedback in order to gauge overall customer satisfaction and identify areas for improvement.

By applying machine learning algorithms and natural language processing techniques to review data, sentiment analysis can classify reviews as positive, negative, or neutral. This can provide valuable insights into customer sentiment towards specific products or aspects of the shopping experience, such as shipping times, customer service, or product quality.

Analyzing customer sentiment in this way can help e-commerce websites to make data-driven decisions about product development, marketing strategies, and customer service initiatives. By understanding the sentiment of their customers, e-commerce websites can tailor their offerings to meet customer needs and increase overall customer satisfaction, leading to greater success in the highly competitive online marketplace.

II. LITERATURE SURVEY

1.Mining and summarizing customer reviews:

The system examines all the reviews on the finished products by all the different customers. Only check the counterfeit features that customers have expressed on the product. Reviews can be positive and sometimes negative. The system works in three steps: First, the characteristics of the finished product reviewed by the customer must be mine. The estimated judgments are then identified and decided which opinion judgment is positive or negative. The last step is the synthesis of the results. M. Hu and B. Liu.

2.Cross-domain co-extraction of sentiment and topic lexicons:

The point of this system is that in the previous works, the stock market supervision method is the best. A description of the mechanism of physically named fact monitoring methods. In this theme, the system includes specified editing limits and we don't need named records either. Unlike the system which requires a lot of tag information. In the first stage, the

Copyright to IJARSCT

www.ijarsct.co.in

DOI: 10.48175/568

280

ISSN
2581-9429
IJARSCT

system generates a small number of high-confidence opinions and case seeds in the target area of influence. In a later step, the system recommends works of fiction. Relationship Bootstrap algorithm. The research findings convey the idea that building the scope of influence system can extract the exact lexicon in the target province. Li, S.J. Pan, O. Jin Qingyang, et al[2]

3. Extracting and ranking product features in opinion documents:
This document focuses on drawing functions. To solve this problem, double replication was introduced. This mechanism is suitable for medium-sized departments. Representing a large and small corpus, it can give low precision and low recall. To address these two issues, two improvements are introduced to strengthen the call for ideas. To get better accuracy for both candidates, feature locations are useful for extracting feature candidates. For candidate status indicators the quality value is determined by two factors: the meaning of the quality and the appearance of the attribute. The crisis is formulated in the form of a bipartite graph and the main algorithm for ranking web pages HITS. Experiments on the dataset give insights into the potential results presented. IZhang, Liu Bao, et al [3]

[4] Mining opinion features in customer reviews:
The system is designed to review all buyer review points for finished products. This summary project is very different from traditional content summaries. The system does not merge revisions by choosing to separate the original sentence from the revision. In this article, only splice is concerned with the removal of features from the opinion pieces examined by reviewers. Flag numbers can be used to provide certain functions.The results indicate that these technologies are of great value. M. Hu and B. Liu [4]

III. PROBLEM DEFINITION

The problem that we aim to address in this study is how to effectively analyze customer feedback and reviews on an e-commerce website in order to understand customer sentiment towards specific products and aspects of the shopping experience. While e-commerce websites typically collect a large volume of customer reviews, it can be difficult to manually analyze and interpret this data in a meaningful way. Sentiment analysis provides a solution to this problem by automating the process of analyzing text data to classify reviews as positive, negative, or neutral.

However, there are several challenges associated with performing sentiment analysis on e-commerce reviews. One challenge is the presence of sarcasm and irony in text, which can be difficult to detect using traditional sentiment analysis techniques. Additionally, reviews may contain misspellings, grammatical errors, and slang, which can impact the accuracy of sentiment analysis algorithms. Finally, there is the challenge of scaling sentiment analysis to handle large volumes of data in real-time, as e-commerce websites generate a continuous stream of customer feedback that must be analyzed in a timely manner.

By addressing these challenges, we aim to develop an effective sentiment analysis approach for e-commerce websites that can provide valuable insights into customer sentiment and opinion, ultimately leading to increased customer satisfaction and success in the online marketplace.
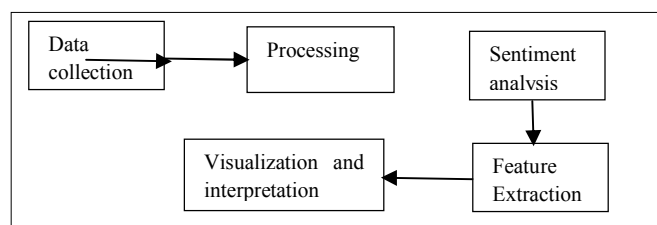
IV. METHODOLOGY

Fig.1: flow of execution

As shown in fig.1 The methodology for sentiment analysis of e-commerce website based on review mining can be divided into the following steps:

- Data collection: The first step is to collect the data in the form of customer reviews from the e-commerce website. This can be done using web scraping techniques or by using APIs provided by the website.
- Preprocessing: Once the data is collected, it needs to be preprocessed to remove noise and irrelevant information. This includes tasks such as removing HTML tags, punctuation, stop words, and converting the text to lowercase.
- Sentiment analysis: The next step is to perform sentiment analysis on the preprocessed data. This involves using machine learning algorithms and natural language processing techniques to classify reviews as positive, negative, or neutral.
- Feature extraction: In addition to sentiment classification, it is also important to extract features from the reviews, such as the most commonly occurring words or phrases within each sentiment category. This can provide further insights into customer sentiment and opinion.
- Visualization and interpretation: The final step is to visualize and interpret the results of the sentiment analysis and feature extraction. This can be done using tools such as data visualization libraries, and can help to identify trends and patterns in customer feedback, as well as areas for improvement in the shopping experience.

V. SENTIMENT ANALYSIS

Sentiment analysis for product reviews is a technique that involves using natural language processing and machine learning algorithms to analyze text data, such as customer reviews, in order to determine the sentiment or emotional tone expressed towards a particular product. The sentiment can be classified as positive, negative, or neutral based on the language used in the review.

Sentiment analysis for product reviews is becoming increasingly important for e-commerce websites and retailers as it can provide valuable insights into customer sentiment towards specific products, features, or aspects of the shopping experience. By analyzing product reviews and identifying the sentiment expressed within them, e-commerce websites can gain a better understanding of customer needs and preferences, identify areas for improvement, and make data-driven decisions to improve customer satisfaction.

Sentiment analysis can also be used to identify trends and patterns in customer feedback, such as recurring complaints or issues, which can help retailers to take corrective action to address these concerns. Overall, sentiment analysis for product reviews is a powerful tool for e-commerce websites to better understand customer sentiment and improve the shopping experience for their customers.

VI. PARTIAL WORD ALIGNMENT MODEL

Partial word alignment is a technique used in sentiment analysis for reviews to improve the accuracy of identifying the sentiment of the review. The idea behind partial word alignment is to consider not only individual words in the review but also partial matches between words and sentiment-bearing phrases.

It follows the following methodology:

1. Identifying sentiment-bearing phrases: The first step is to identify the sentiment-bearing phrases in the review. These are phrases that carry a positive or negative sentiment, such as "not bad," "very good," or "terrible." There are different ways to identify sentiment-bearing phrases, such as using a pre-defined list of phrases or using machine learning techniques to extract phrases from the reviews.

2. Aligning words with sentiment-bearing phrases: Once the sentiment-bearing phrases are identified, the next step is to align individual words in the review with these phrases. The alignment process involves matching the words with the sentiment-bearing phrases based on their position and context in the review. For example, if the phrase "not bad" appears in the review, the model would align the word "good" with the phrase based on its position and context.

3. Assigning sentiment scores to words: After the alignment process, the model assigns sentiment scores to individual words based on their alignment with sentiment-bearing phrases. Words that are aligned with positive phrases, such as "very good" or "excellent," are assigned a positive sentiment score, while words that are aligned with negative phrases, such as "terrible" or "awful," are assigned a negative sentiment score.

4. Combining sentiment scores: The final step is to combine the sentiment scores of individual words to determine the overall sentiment of the review. There are different ways to combine the scores, such as using a simple sum or a weighted average of the scores. The resulting sentiment score can be classified as positive, negative, or neutral based on a threshold value.

VII. FEATURE BASED ANALYSIS

Features based sentiment analysis of product reviews is a technique that involves identifying the specific product features that are mentioned in a review and analyzing the sentiment associated with each feature. The first step is to extract the features mentioned in the review. This can be done using techniques such as named entity recognition or pattern matching. For each feature mentioned in the review, the sentiment associated with that feature is analyzed. The sentiment analysis may be performed on the feature itself, or on the words that are associated with the feature in the review. The sentiment scores of the features are aggregated to compute an overall sentiment score for the product. The features-based sentiment analysis model is trained on a labelled dataset of reviews, where each review is labelled as positive, negative, or neutral. The features-based sentiment analysis technique has several advantages over other sentiment analysis techniques, as it allows for a more fine-grained analysis of the sentiment associated with specific product features. This can be useful for product development and marketing purposes, as it allows for a better understanding of the strengths and weaknesses of a product.

VIII. FEATURE WISE GRAPH GENERATION

Feature wise graph generation from data collected through sentiment analysis based on product reviews involves creating a graph that displays the sentiment associated with specific product features
It works through following phases.

1. Data collection: The first step is to collect a dataset of product reviews and perform sentiment analysis on each review. The sentiment analysis results should include the sentiment score for each feature mentioned in the review.
2. Feature grouping: The next step is to group the features mentioned in the reviews into categories. This can be done manually or using techniques such as clustering or topic modelling.
3. Graph generation: For each category of features, a graph is generated that displays the sentiment associated with each feature over time. The sentiment score can be represented using various visualization techniques, such as a line chart or a bar chart. The x-axis of the graph represents time, while the y-axis represents the sentiment score.

The generated graph can be used to identify trends and patterns in the sentiment associated with specific product features over time. This can help in identifying areas of the product that need improvement or in understanding the impact of changes made to the product.
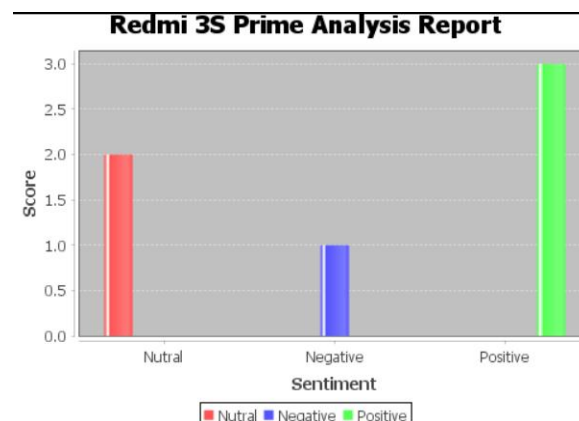
Some Result of
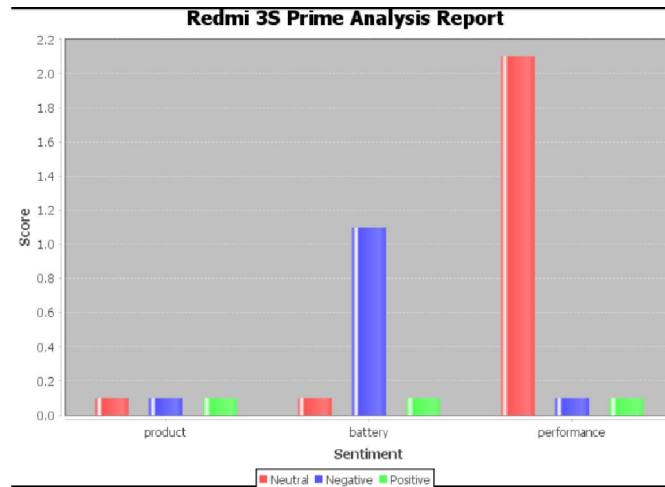

Fig 7.1: Sentiment Based Graph

Fig 7.2: Feature Based Graph

IX. CONCLUSION

In this paper, we describe the work of extracting words using a partially supervised word alignment model. The goal is to extract opinion words and opinion objects and to detect opinion relationships between them using a partially supervised word alignment model. The dynamic contribution focuses on sentiment analysis of customer reviews and classifies them into positive reviews, negative reviews and detects the relationship between opinion targets and opinion words. Here the model gives positive and negative product reviews so that customers can decide whether to buy product and the manufacturer idea to improve the quality of product over time. Sentiment Polarity Detection can produce more accurate mining results than any other leading systems.

REFERENCES

[1]. Kang Liu,Liheng Xu, and Jun Zhao, "Co-extracting opinion targets and opinion words from online reviews base on the word alignment model" march 2015.

[2]. M. Hu and B. Liu, Mining and summarizing customer reviews, in Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Seattle, WA, USA, 2004, pp. 168177.

[3]. F. Li, S. J. Pan, O. Jin, Q. Yang, and X. Zhu, Cross-domain coextraction of sentiment and topic lexicons, in Proc. 50th Annu. Meeting Assoc. Comput. Linguistics, Jeju, Korea, 2012, pp. 410419.

[4]. L. Zhang, B. Liu, S. H. Lim, and E. OBrien-Strain, Extracting and ranking product features in opinion documents, in Proc. 23th Int. Conf. Comput. Linguistics, Beijing, China, 2010, pp. 1462 1470.

[5]. M. Hu and B. Liu, Mining opinion features in customer reviews, in Proc. 19th Nat. Conf. Ar tif.Intell., San Jose, CA, USA, 2004, pp. 755760.