

A Data-Driven Approach for Understanding Customer Sentiment Trends in E-Commerce

¹ Dr.V.Venkateswara Rao,² Harshitha Mutta,³ Cherish Vuppala,⁴ Sathvika Ganapathineedi,⁵ Kvs Mounika,⁶ MUTTA KARTHIK PATTABHI RAMA RAO,⁷ Meghana Manda

¹ Professor, ^{2,3,4,5,6,7} Final BTech

^{1,2,3,4,5,6,7} Dept of CSE, Sri Vasavi Engineering College, Tadepalligudem

Abstract

In the rapidly evolving backdrop of e-commerce, customer feedback plays a crucial role in shaping the accomplishment or disappointment of a business. Using the Valence Aware Dictionary and Sentiment Reasoner, also known as the VADER library, we can gain insights from customer reviews. Customer reviews provide a rich source of information that allows us to measure the prevailing sentiments among consumers. By employing this sophisticated tool, our research aims to offer prized understandings into the sentiments expressed by customers regarding various products or services. As the e-commerce industry thrives on customer satisfaction, understanding sentiment trends through data-driven approaches becomes increasingly vital for staying competitive and ensuring sustained success in the marketplace.

Key Words: E commerce, Feedback, Sentiment Analysis, Machine Learning, Data Analytics

Introduction

Customer input plays a critical role in determining a business's future in the ever-changing world of e-commerce. The present investigation utilizes the Valence Aware Dictionary and Sentiment Reasoner, which is also known as the VADER library, to explore the complex terrain of customer evaluations. Given the critical role that consumer sentiments play in determining a company's success or failure, our research aims to extract insightful information from the large body of customer feedback. We want to identify and examine consumer sentiments around various products and services by utilizing the advanced analytical tools provided by the VADER library.

Our project addresses the time-consuming nature of reviewing all the customer feedback by implementing an efficient solution. Recognizing the contest of reading numerous reviews, our approach aims to save valuable time for businesses, offering a streamlined process for extracting insights from customer sentiments in the fast-paced world of e-commerce.

Our proposed solution involves presenting a visualized overview of analyzed reviews. By transforming data into a visual format, businesses can swiftly grasp insights from customer sentiments. This approach simplifies the review analysis process, enhancing efficiency and providing a quick, comprehensive view of customer feedback in the e-commerce landscape. The ability of sentiment analysis to extract significant insights from huge volumes of textual data has recently sparked the interest of many in natural language processing. A customer assessment or social media stake, for example, could be the subject of a sentiment analysis, which classifies the sentiment, attitude, and opinion expressed in the text as positive, negative, or neutral. Businesses seek this data because it can support them enhanced apprehend their consumers' perspectives and improve the quality of their goods and services.

Literature Survey

Sentiment classification traditionally involves two approaches: supervised and unsupervised, focusing on semantics. The efficiency of these methods relies heavily on accurately extracting the set of features essential for sentiment detection.

In 2012, Mudinas established the scheme in which vocabulary and learning-based method pooled for perception level sentiment analysis. Vinodhini offered a study that concealed the approaches and practices in sentiment analysis as well as experiments that acted in the arena. They paralleled the several methods for sentiment classification. The methods associated were machine learning methods by means of supervised and unsupervised with dissimilar procedures and the feature based sentiment classification etc. They presented a graphical result viewing a appraisal of these methods.

In 2013, Hemalatha established a system in which pre-processing and machine learning techniques pooled to accumulate tweets from social networking situates. The noise in the data was disinterested by means of **preprocessing** methods. Machine learning techniques were functional on those tweets which progress business brainpower by providing some forecast for decision making. Taxonomy of results of definite issue analysis was done as Positive, Negative and Neutral.

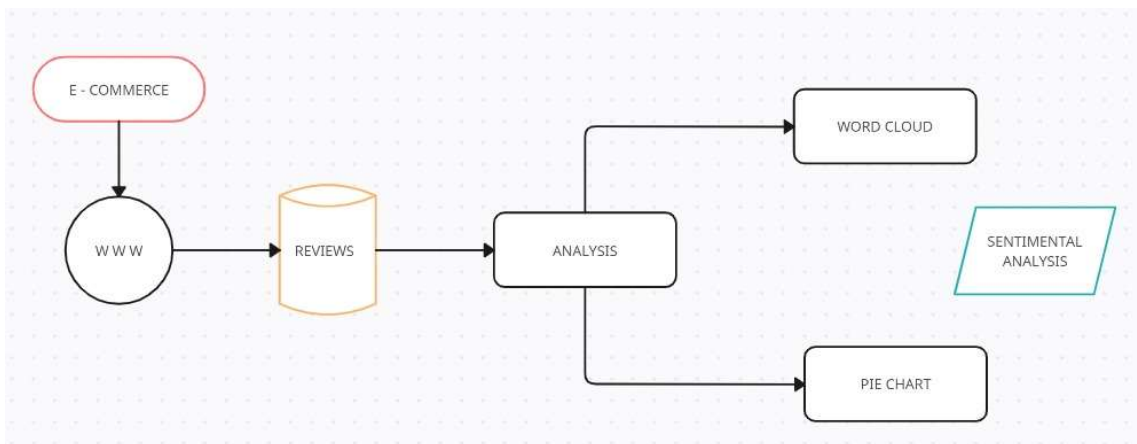
In 2014, Pimpalkar established a system that displays the remarks and feedbacks/reviews for products. They determined the polarity of sentiments for the comments of the person. After this the contrast between two different products was done using comments which were acknowledged from the online resources. This judgement leads to find the greatest product.

Motivation

The motivation behind our project is understanding and reading the sentiments expressed by customers to provide precious comprehensions into their preferences, concerns, and overall satisfaction levels. By delving into sentiment analysis, our project seeks to empower e-commerce businesses to swiftly comprehend and respond to customer feedback. This pre-emptive method not only improves customer relations but also contributes to strategic decision-making, ultimately adopting a more competitive and customer-centric environment in the e-commerce industry.

In addition, motivation extends to recognition of the time-consuming nature of reading and understanding each customer's review manually. Using tools such as the Valence Aware Dictionary and Sentiment Reasoner (VADER) to integrate feelings analysis in projects, we aim to offer businesses with practical and efficient ways to refine customer's feelings. This motivation is based on the understanding that visual representation of major tendencies can not only save valuable time for companies but also enable them to make data motivated decisions and ensure competitive advantages in a dynamic and competitive e-commerce environment.

System Model



Problem Definition

The problem we're tackling in our project is that it's hard for businesses to rapidly recognize what consumers are saying in a large number of online reviews. Reading all these reviews takes a lot of time and effort. So, we aim to solve this by creating a visual way for businesses and customers to see and grasp the sentiments in

customer reviews easily, helping them save time and make better decisions in the fast-paced world of E-commerce.

Proposed Work

In our proposed project, we aim to conduct sentiment analysis on customer reviews using the VADER (Valence Aware Dictionary and Sentiment Reasoner) tool. Sentiment analysis plays a critical role in understanding customer opinions and sentiments towards products or services. VADER, a pre-built sentiment analysis tool, is well-suited for this task as it is specifically designed to handle social media text and provides a rich understanding of sentiments, including the intensity of positive and negative expressions. The first step of our project involves collecting customer reviews from various online platforms related to the target product or service. Once the data is gathered, we will preprocess it by removing noise, handling missing values, and tokenizing the text to prepare it for analysis. During the preprocessing phase, we aim to clean the data and structure it in a way that facilitates meaningful sentiment analysis.

The next phase involves implementing the VADER sentiment analysis tool. VADER utilizes a combination of lexical and grammatical heuristics, along with a sentiment lexicon, to analyze text and determine the polarity of sentiments. The tool provides a polarity score for each review, indicating the level of positivity, negativity, or neutrality. The compound score, a key metric in VADER, is calculated established on the quantity of individual polarity scores and a normalization factor.

The compound score is calculated using the formula:

$$\text{Compound Score} = \frac{\sum \text{Polarity Scores}}{\sqrt{\sum \text{Polarity Scores}^2 + \text{Normalization Constant}}}$$

Where:

- $\sum \text{Polarity Scores}$ is the sum of the individual polarity scores of each word in the sentence.
- $\sqrt{\sum \text{Polarity Scores}^2 + \text{Normalization Constant}}$ is a normalization factor to prevent bias toward longer sentences.

We plan to leverage these polarity scores to classify the reviews into positive, negative, or neutral categories. Additionally, to enhance the interpretability of the results, we intend to visualize the outcomes using pie charts and word clouds. These visualizations will offer a clear understanding of the distribution of sentiments within the customer reviews.

we can use a threshold to categorize the compound scores into different sentiment categories (positive, negative, neutral). Here's an example of how you might implement this in the context of your project:

Let's say you decide on the following threshold values:

- If Compound Score ≥ 0.05 , classify as Positive.
- If $-0.05 < \text{Compound Score} < 0.05$, classify as Neutral.
- If Compound Score ≤ -0.05 , classify as Negative.

The project's significance lies in its potential to provide businesses with valuable insights into customer perceptions and preferences. By comprehending the sentiments expressed in customer reviews, companies can make informed decisions to improve their products or services, enhance customer satisfaction, and ultimately strengthen their market position. This approach not only benefits businesses but also aids customers in making informed decisions, saving time, and ensuring a more satisfactory consumer experience.

Overall, our proposed work on sentiment analysis using VADER aims to contribute to the field of customer feedback analysis, providing a practical and efficient solution to gain actionable insights from customer reviews. Through a systematic application of sentiment analysis techniques, we aim to create a valuable resource for businesses seeking to understand and respond to customer sentiment effectively.

Conclusion

In conclusion, our sentiment analysis using the Vader algorithm effectively processed customer reviews, generating insightful visual representations. By carefully examining the sentiments expressed in the reviews, we obtained a comprehensive overview of customer opinions. The visualizations offer a user-friendly snapshot, making it easier for commerce to interpret and act upon the feedback. This survey paper highlights the importance of leveraging Vader for sentiment analysis, emphasizing its practicality in enhancing decision-making processes based on customer reviews. Overall, our approach provides a valuable tool for businesses seeking actionable insights from customer sentiments, ultimately contributing to improved products and services.