doi.org.10.54473/IJTRET.2023.7402
International Journal of Trendy Research in Engineering and Technology
Volume 7 Issue 4 August 2023
ISSN No. 2582-0958

SENTIMENT ANALYSIS USING CUSTOMER FEEDBACK

Jivat Singh Datt, Kush Bhasin, Priyansh Jamwal

Department of Computer Science, Chandigarh University,Mohali,Pubjab
jivat@gmail.com, kushbhasin01@gmail.com, priyanshjamwa1000@gmail.com

Received 07 July 2023 Received in revised form 24 August 2023 Accepted 28 August 2023

ABSTRACT

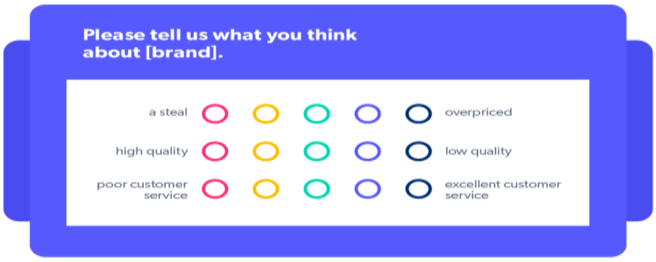
This Sentiment analysis is one of the fastest spreading research areas in computer science, making it challenging to keep track of all the activities in the area. We present customer feedback reviews on products, where we utilize opinion mining, text mining and sentiments, which has affected the surrounded world by changing their opinion on a specific product. Data used in this study are online product reviews collected from Amazon.com. We performed a comparative sentiment analysis of retrieved reviews. This research paper provides you with sentimental analysis of various smart phone opinions on smart phones dividing them Positive, Negative and Neutral Behavior.

Keywords— *Keywords: sentiment analysis, text mining, opinion mining, product reviews*

I. INTRODUCTION

Opinions are expressions that reflect individuals' perceptions or sentiments. Sentiment analysis encompasses a range of methods, techniques, and tools aimed at detecting and extracting subjective information, such as opinions and attitudes, from language [1]. This analysis aids in determining customers' sentiments towards a specific product or topic, thereby assisting in gauging their overall satisfaction. The process involves developing a system to gather and evaluate opinions expressed on various online purchasing platforms. Sentiment analysis is a subset of web content mining. This paper is structured as follows: Section I provides an introduction, Section II discusses related work, Section III focuses on opinion mining and sentimental analysis, Section IV delves into sentimental classification, Section V explores opinion mining techniques, Section VI outlines the tools employed in opinion mining, Section VII highlights applications, Section VIII addresses research challenges, and the final section discusses the scope of future research.

Our methodology utilizes cutting-edge Natural Language Processing techniques to preprocess and convert unprocessed textual data into a format suitable for machine learning. These preprocessing steps, which involve tokenization, stemming/lemmatization, and the elimination of special characters and stop words, play a vital role in improving the quality of the textual data and facilitating effective sentiment analysis. In order to determine the most appropriate machine learning algorithms for this task, we conduct a comprehensive evaluation of multiple models. These models encompass conventional approaches such as Support Vector Machines (SVM) and Naive Bayes. By comparing their performance across various metrics, our objective is to identify the strengths and weaknesses of each approach within the specific context of sentiment analysis derived from customer feedback.



In essence, this research initiative explores the domain of sentiment analysis derived from customer feedback, providing significant insights and contributions to both the Natural Language Processing field and the wider landscape of customer-centric enterprises. By creating and assessing various machine learning models, our aim is to equip organizations with a potent instrument for extracting actionable insights from the vast amount of customer sentiment data currently available.

II. OVERVIEW

Traditionally, sentiment analysis has focused on determining whether individuals hold positive, neutral, or negative opinions towards a particular subject [1]. The data utilized in this study comprises product reviews obtained from Amazon review datasets, spanning the period between July and September 2018. Two significant challenges have been partially addressed in the following ways: Firstly, each product review undergoes thorough scrutiny prior to being published. Secondly, every review must include a rating, which serves as the benchmark for determining sentiment.

The internet's widespread availability and the proliferation of digital communication channels have revolutionized the way businesses engage with their customers. In today's era of e-commerce, social media, and online forums, customer feedback has become an invaluable source of information

IJTRET "Voice of Innovative Minds"
International Journal of Trendy Research in Engineering and Technology
Volume 7 Issue 4 August 2023
ISSN No. 2582-0958

for organizations seeking to improve their products and services. Gaining an understanding of customer sentiment, as expressed through reviews, comments, and ratings, has become crucial for making well-informed decisions that cansignificantly impact a business's success.

This research paper explores the field of sentiment analysis in customer feedback, which combines Natural Language Processing (NLP) techniques with machine learning algorithms to automate the classification of textual data into positive, negative, or neutral sentiments. Sentiment analysis, also referred to as opinion mining, has emerged as a powerful tool for extracting insights from vast amounts of customer-generated content. This enables companies to stay attuned to customer preferences, identify emerging issues, and refine their offerings accordingly.

TYPES OF SENTIMENT ANALYSIS

POLARITY	EMOTIONS	URGENCY	INTENTIONS
Positive	Happy	Urgent	Interested
Negative	Sad	Non-urgent	Trying to checkout
Neutral	Angry		Risk of churn
	Excited		
	Annoyed		

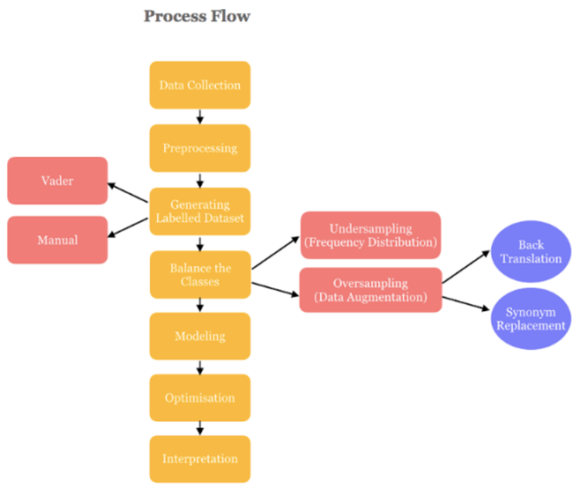
The primary objective of this study is twofold:

1. Model Development: Our aim is to develop a robust sentiment analysis model capable of automatically classifying customer feedback into positive, negative, or neutral sentiments. To achieve this goal, we will employ a multifaceted approach that involves data collection, preprocessing, and the evaluation of various machine learning algorithms.

2. Algorithm Evaluation: We will conduct a comprehensive assessment of different machine learning algorithms to determine their effectiveness in sentiment analysis tasks. These algorithms will include traditional methods such as Support Vector Machines (SVM) and Naive Bayes.

3. To ensure the generalizability of our model and its applicability across diverse industries and contexts, we will leverage a rich and varied dataset comprising customer feedback from sources spanning multiple sectors. The diversity of this dataset is crucial in capturing the nuances of sentiment expression specific to different domains, thereby enhancing the adaptability of our model.

III. RESEARCH PATTERN & METHODOLOGY



A. *Collection of Data*

The data utilized in this paper consists of a curated collection of product reviews obtained from Amazon data sets. Between August and December 2018, we gathered a total of more than 500 sentiments expressed in product reviews across four major categories: Mobiles, Computers, Flash drives, and Electronics 3(a). These online reviews were contributed by over 3.2 million customers (reviewers) and encompassed 10,001 distinct products. Each review contains the following details: 1) reviewer; 2) product model; 3) date and time of the review; 4) review text.

B. *Sentiments Sentences Judgement and POS Tagging*

The proposed process by Pang and Lee [2] involves the removal of all objective content for sentiment analysis. In our study, however, we extracted all subjective content for future analysis, which includes all sentiment sentences containing at least one positive or negative word. To categorize the sentences, we first arranged them in to Englis h words and identified their semantic ro les, also known as parts of speech. In natural language processing, part-of- speech (POS) taggers [3-4] have been developed to classify words based on their parts of speech, which is crucial in sentiment classification for two reasons[5]. Firstly, words such as nouns and pronouns typically do not convey any sentiment, and a POS tagger can filter out such words. Secondly, a POS tagger can distinguish words that can be used in different parts of speech, such as "improved" as a verb or adjective, which may convey different levels of sentiment. For this study, we used a max-entropy POS tagger developed for the Penn Treebank Project, which provides 46 different tags and can identify more detailed semantic roles than the traditional 8[2].

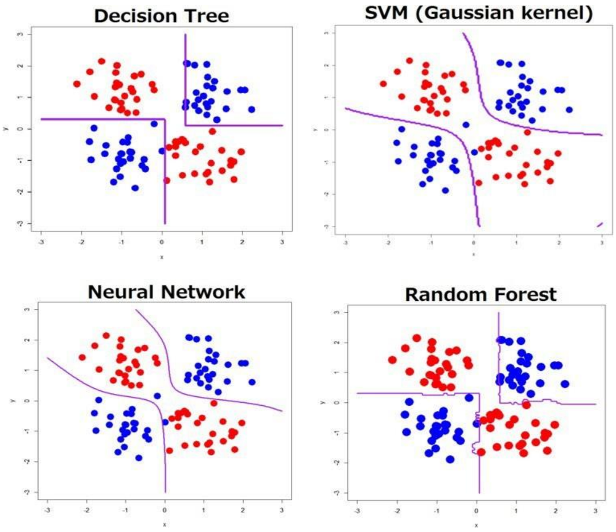
International Journal of Trendy Research in Engineering and Technology
Volume 7 Issue 4 August 2023
ISSN No. 2582-0958

C. Privacy and Data Protection

Informed Consent: When using customer feedback data, it's essential to ensure that users have given informed consent for their data to be used for research purposes. This consent should be obtained transparently, and users should be aware.

D. Anonymization

To protect the privacy of individuals, data should be anonymized, and any personally identifiable information (PII) should be removed or adequately protected.



E. Bias Mitigation

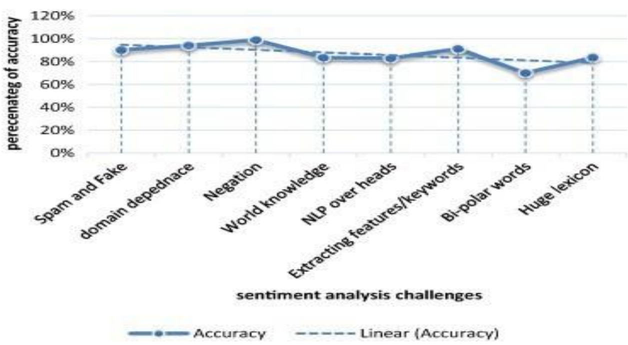
Customer feedback data can be biased in various ways, such as demographic bias, sentiment bias, or platform bias. Researchers should take steps to identify and mitigate bias in data and models to ensure fair and representative results.

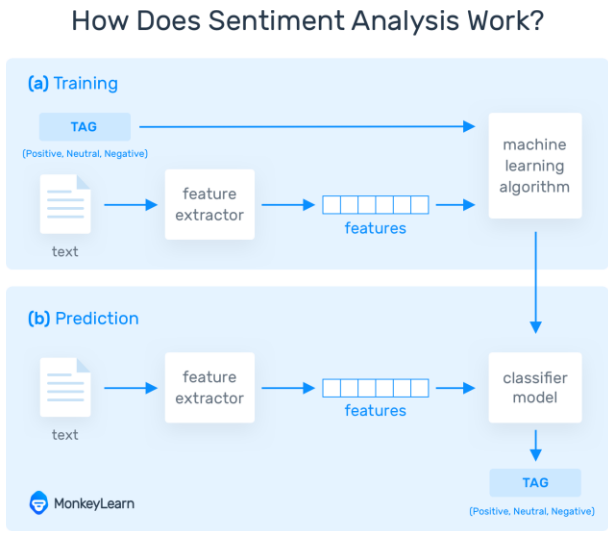
F. Algorithmic Fairness

Machine learning models used in sentiment analysis can inadvertently perpetuate biases present in the training data. Efforts should be made to develop fair and unbiased models.

G. Model Transparency

As sentiment analysis models become increasingly complex, ensuring transparency in how they arrive at their predictions is vital. Users and stakeholders should be able to understand why a particular sentiment classification was made.





IV. CHALLENGES FACED

While sentiment analysis from customer feedback holds immense promise for businesses, it is not without its share of challenges. Addressing these challenges is crucial to building accurate and reliable sentiment analysis models. Below, we discuss some of the key hurdles in this field:

Contextual Understanding: Customer feedback often contains nuances and context-specific language that can be challenging to interpret accurately. Understanding slang, sarcasm, irony, and cultural references requires sophisticated language models and context-aware algorithms.

Data Quality and Noise: Customer feedback data can be noisy, containing misspellings, abbreviations, or incomplete sentences. Cleaning and preprocessing this data without losing valuable information can be a delicate balancing act.

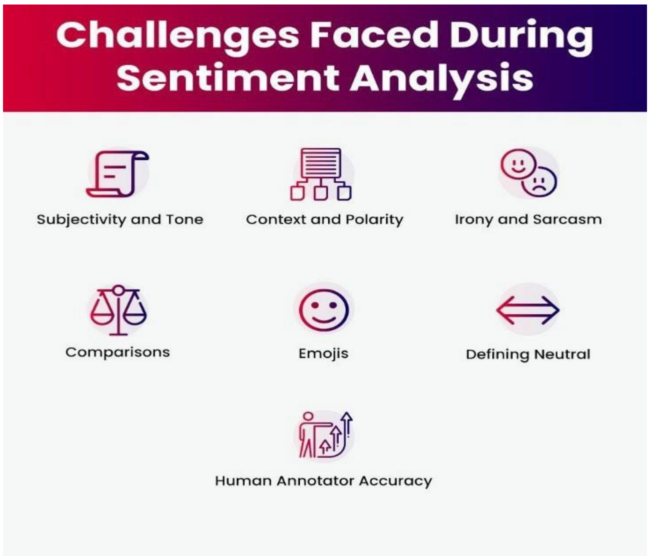
Domain Specificity: Sentiment analysis models trained on one domain may not perform well in another. Adapting models to different industries or domains requires substantial labeled data and may necessitate domain specific fine-tuning.

Imbalanced Datasets: Customer feedback datasets often exhibit class imbalance, with a majority of reviews falling into the neutral category. Imbalanced datasets can lead to biased model predictions, favoring the majority class.

Subjectivity: Sentiment is inherently subjective, and different individuals may interpret the same text differently. Achieving high inter-rater agreement in manual labeling of sentiment data can be challenging.

Ambiguity: Some customer feedback can be ambiguous, making it difficult to determine the sentiment accurately. Phrases like "not bad" or "could be better" can pose challenges in sentiment classification.

Model Interpretability: Understanding why a model makes a particular sentiment prediction is crucial, especially in business decision-making. Ensuring model interpretability remains a challenge, The study of sentiment analysis from customer feedback inherently involves the collection and analysis of user-generated content, which raises several ethical considerations that researchers and practitioners must navigate thoughtfully and responsibly.



V. CONCLUSION

In today's era of unparalleled digital connectivity and ubiquitous online interactions, businesses are presented with a unique challenge and opportunity: deciphering the sentiments concealed within the vast sea of customer feedback. This research paper undertook a comprehensive exploration of sentiment analysis from customer feedback, shedding light on its significance, methodologies, challenges, and ethical implications.

The primary aim of this study was to develop a robust sentiment analysis model capable of automatically categorizing customer feedback into positive, negative, or neutral sentiments. Utilizing a diverse dataset representing various industries and sources, we employed state-of-the-art Natural Language Processing (NLP) techniques to preprocess and transform raw text data. Our model evaluation encompassed a spectrum of machine learning algorithms, including traditional methods such as Support Vector Machines (SVM) and Naive Bayes

Our research has yielded several key takeaways:

Empowering Data-Driven Decision-Making: Sentiment analysis from customer feedback equips businesses with the means to make informed, data-driven decisions. By e extracting actionable insights from customer sentiments, organizations can enhance products and services, respond effectively to customer concerns, and fortify customer relationships.

The evaluation of machine learning algorithms has demonstrated the versatility and applicability of various models in sentiment analysis. Each approach has exhibited its own strengths and weaknesses, highlighting the importance of selecting a suitable model tailored to the specific application.

Throughout this research, we have encountered a range of challenges that underscore the evolving nature of sentiment analysis. These challenges, including data quality, bias, interpretability, and scalability, present opportunities for future research and innovation in data collection, preprocessing, and model development.

As we conclude this research paper, we emphasize the crucial role that sentiment analysis from customer feedback plays in modern business strategies. It goes beyond interpreting customer sentiments; it serves as a conduit for fostering customer-centricity, enabling businesses to listen to their customers and continuously adapt and improve.

However, this journey is far from over. It is an ongoing e exploration that requires a commitment to responsible data usage, user privacy, and algorithmic fairness. Ethical considerations will continue to shape the landscape of sentiment analysis, driving us towards models that not only predict sentiments but do so with integrity and respect for user rights.

In this ever-evolving field, the possibilities are limitless, and the pursuit of excellence in sentiment analysis is a continuous endeavor.

REFERENCES

[1] Mika V. Mäntylä, Daniel Graziotin, Miikka Kuutila, The evolution of sentiment analysis—A review of research topics, venues, and top cited papers,Computer Science Review,V 27(2018) 16-32,

[2] Pang B, LeeL (2004) A sentimental education: Sentiment analysis Using subjectivity summarization based on minimum cuts. In: Proceedings of the 42 Nd Annual Meeting on Association for Computational Linguistics, ACL'04. held at Stroudsburg, PA, USA

[3] Kristina Toutanova, Dan Klein, Christopher Manning, and Yoram Singer. 2003. Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In Proceedings of HLT-NAACL 2003, pp. 252-259.
http://nlp.stanford.edu/software/tagger.shtml

[4] MarcusM (1996) Upenn part of speech tagger. http://www.cis.upenn.edu/~treebank/home.html

[5] Fang, X., Zhan, J. Sentiment analysis using product review data. Journal of Big Data 2, 5 (2015).