Sarcasm Detection in Product Reviews for Sentiment Analysis

Bachelor of Science in Information Technology
(Mobile Application and Information Security)

Under

Assam Down Town University

Submitted by

Student1
Dijumoni Ray
ADTU/2020-23/BIT/002
dijumoniray@gmail.com
9127355251
6th Semester, BSc-IT(MAIS)

Student2
Sudhir Kumar Gupta
ADTU/2020-23/BIT/009
Sudhirkrg494@gmail.com
8638696675
6th Semester, BSc-IT(MAIS)

Student3
Fahim Ahmed
ADTU/2020-23/BIT/007
fahim9365100@gmail.com
9365666538
6th Semester, BSc-IT(MAIS)

Department: Faculty of Computer & Technology

Under the Guidance of

Mr. Mehboob Hassan Ben Siraj Barbhuyan
Assistant Professor, Computer & Te
Faculty of Computer & Technology
Assam Down Town University,
Guwahati-26, Assam

INTRODUCTION

Sentiment analysis, also known as opinion mining, is a natural language processing technique. The goal of sentiment analysis is to determine the sentiment, or emotional tone, of a piece of text, such as a tweet, product review, or news article.

Sentiment analysis on sarcasm detection in product reviews is a process of using natural language processing and machine learning techniques to identify sarcastic comments in customer reviews. The purpose of this analysis is to detect and extract sarcastic comments in product reviews, which can provide valuable insights into customer sentiment towards the product.

Sarcasm detection can be particularly useful for companies that want to understand how customers feel about their products. By identifying sarcastic comments, companies can gain insights into areas where their products may be falling short and take steps to address those issues.

Sarcasm detection in product reviews can also be useful to customers by identifying sarcastic comments in product reviews, customer can gain a more accurate understanding of the overall sentiment towards the product and make more informed decisions about whether to purchase it or not.

In this project, we propose using sarcasm detection in product reviews using sentiment analysis. The goal of this approach is to improve the accuracy and usefulness of product review analysis by identifying sarcastic statements. This can provide businesses with valuable insights into customer sentiment and help guide their decision-making processes.

FEASIBILITY STUDY

- **Data availability**: The data we will be using to train and test the model are available therefore it is feasible.
- **Technology requirements:** Sentiment analysis requires natural language processing (NLP) techniques and which are available or can be acquired within budget constraints therefore it is feasible.
- **Cost-benefit analysis**: The potential costs and benefits of implementing sentiment analysis should include the costs of acquiring and maintaining technology therefore it is feasible.
- **Accuracy and reliability**: The accuracy and reliability of sentiment analysis algorithms should be assessed. This can be done by testing the algorithms on a sample of data and comparing the results with human evaluations.
- **Legal and ethical considerations**: Sentiment analysis may raise legal and ethical concerns, particularly if it involves analysing personal data.We will ensure that appropriate measures are in place to protect privacy and data security therefore it is feasible.

OBJECTIVES

- Sarcasm can significantly impact the sentiment of a product review. By detecting sarcasm, we can better understand the true sentiment of a review and improve the accuracy of sentiment analysis.

- Sarcasm can often be used to express dissatisfaction or frustration in a humorous or ironic manner. By detecting sarcasm, businesses can gain valuable insights into customer sentiment and identify areas where improvements can be made.

- Detecting sarcasm manually can be time-consuming and subjective. By automating the process using machine learning-based approaches, we can improve the efficiency and accuracy of sarcasm detection in product reviews.

- There are various machine learning algorithms and feature extraction methods that can be used for sarcasm detection. By exploring the effectiveness of different approaches, we can identify the most accurate and efficient methods for detecting sarcasm in product reviews.

- Sarcasm can be complex and context-dependent, and there may be cultural and linguistic differences in the use of sarcasm. By exploring the limitations and challenges in sarcasm detection, we can identify areas where improvements can be made and guide future research in this area.

PROBLEM STATEMENT

The problem with analyzing product reviews is that the tone and sentiment of the text can be challenging to interpret, especially when it comes to detecting sarcasm. Sarcasm is a form of irony where the intended meaning is the opposite of the literal meaning of the words used. This makes it difficult for traditional sentiment analysis techniques to accurately detect sarcasm in text. The use of sarcasm in product reviews can significantly impact the sentiment of the review, leading to an inaccurate understanding of customer sentiment. Therefore, there is a need for a more accurate and efficient approach to detect sarcasm in product reviews using sentiment analysis. The problem statement for this research is to explore the feasibility of using machine learning-based approaches to detect sarcasm in product reviews, identify the most effective methods for feature extraction and algorithm selection, and evaluate the accuracy and limitations of the approach. By addressing this problem, businesses can gain a more accurate understanding of customer sentiment and respond appropriately to address any issues or concerns

METHODOLOGY

- **Data Collection**: There are some datasets available on the web if it's not enough for both training and testing the model then We intended to collect the dataset on product reviews with labels indicating the presence or absence of sarcasm.

- **Preprocessing**: In this step we intended to clean the raw text data by removing stop words, punctuations, and other irrelevant information. This step also involves tokenization, stemming, and normalization of the text data.

- **Feature Extraction**: This step involves identifying and extracting the relevant features from the preprocessed text data.

- **Training**: This step involves training the machine learning models using the labeled dataset of product reviews. Various machine learning algorithms such as Naive Bayes, Support Vector Machines (SVM), and Random Forest can be used for training.

- **Evaluation**: We check the performance of the trained machine learning models by comparing their predictions with the actual labels.

- **Deployment**: This step involves integrating the models into the business workflow and using them to analyze new product reviews.

- **Continuous Improvement**: Monitoring the performance of the deployed models and incorporating new data and feedback to improve the accuracy of sarcasm detection over time

EXPECTED OUTCOME

- **Increased accuracy in sentiment analysis**: By detecting sarcasm, the accuracy of sentiment analysis can be significantly improved. This can provide businesses with more accurate insights into customer sentiment and enable them to make informed decisions to improve their products or services.

- **Improved customer satisfaction**: By identifying sarcastic comments in product reviews, businesses can respond appropriately to address any issues or concerns that customers may have. This can improve customer satisfaction and loyalty, leading to increased sales and revenue.

- **Efficient detection of sarcasm**: Automating the process of sarcasm detection using machine learning-based approaches can significantly improve the efficiency of analyzing product reviews. This can save time and resources, enabling businesses to make faster decisions and respond to customer feedback more quickly.

- **Identification of areas for improvemen**t: By analyzing sarcastic comments in product reviews, businesses can identify areas where improvements can be made to their products or services. This can help them to address issues and provide a better overall customer experience.

- **Identification of limitations and challenges**: By exploring the limitations and challenges of sarcasm detection in product reviews using sentiment analysis, future research can be guided towards addressing these issues and improving the accuracy and effectiveness of the approach.

FACILITIES REQUIRED FOR PROPOSED WORK

HARDWARE REQUIREMENTS:

- Processor – Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.70 GHz
- Hard Disk – 1 TB
- SSD – 120GB
- Memory – 8 GB RAM

SOFTWARE REQUIREMENTS:

- Anaconda

BIBLIOGRAPHY

[1] Bhattacharyya, S., Bhattacharya, P., & Kalita, J. K. (2018). A comprehensive survey on fraud detection techniques. ACM Computing Surveys (CSUR), 50(2), 22.

[2]Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2015). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784-3797.

[3] Javvaji Venkataramaiah, Bommavarapu Sushen, Mano. R, Dr. Gladispushpa Rathi, “An enhanced mining leading session algorithm for fraud app detection in mobile applications” International Journal of Scientific Research in Engineering., April2017.

[4] Avayaprathambiha. P, Bharathi. M, Sathiyavani. B, Jayaraj. S “To Detect Fraud Ranking For Mobile Apps Using SVM Classification” International Journal on Recent and Innovation Trends in Computing and Communication, vol. 6, February2018.

[5] Esther Nowroji and Vanitha, Detection Of Fraud Ranking For Mobile App Using IP Address Recognition Technique, vol. 4.

[6] S.R. Srividhya and S. Sangeetha, A Methodology to Detect Fraud Apps Using Sentiment Analysis.

[7] Shashank Bajaj, Nikhil Nigam, Priya Vandana and Srishti Singh, "Detection of fraud apps using sentiment analysis", International Journal of Innovative Science and Research Technology.

[8] Harpreet Kaur, Veenu Mangat and Nidhi, A Survey of Sentiment Analysis techniques.

[9] Dubey Veena and G. D., "Sentiment Analysis Based on Opinion Classification Techniques: A Survey", International Journal of Advanced Research in Computer Science and Software Engineering, pp. 53-58, 2016.

[10] G.D. Upadhye, D.Pise,“Grading of Harvested Mangoes Quality and Maturity Based on Machine Learning Techniques”,IEEE International conference on smart city and Emerging Technology,2018.