© July 2024| IJIRT | Volume 11 Issue 2 | ISSN: 2349-6002

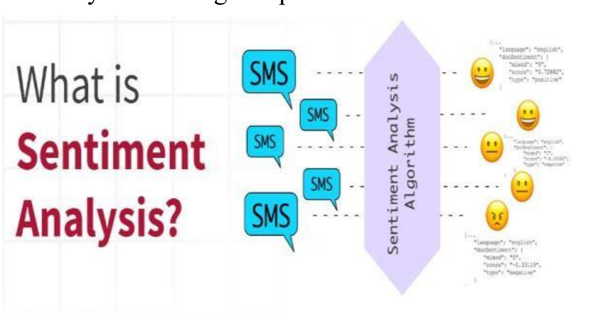
AI Sentimental Analysis Using Customer Feedback

Akash Sharma
Amity School of Engineering and Technology, Amity University Noida, Uttar Pradesh, India

***Abstract* - AI Sentiment analysis of customer feedback assists organizations in understanding client feelings and attitudes, allowing them to enhance products, services, and overall customer satisfaction. Companies can determine total customer sentiment by analyzing text for positive, negative, or neutral sentiments and identifying are as that require work. Sarcasm identification is critical because sarcastic remarks might cause sentiment analysis to produce a false positive or negative result. Accurate sentiment analysis and sarcasm detection give vital information for decision-making and strategic improvements.**

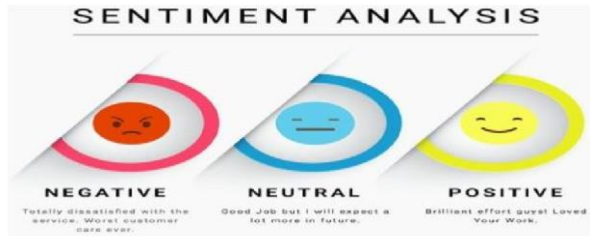
INTRODUCTION

In the current digital environment, companies get a ton of customer input via social media, internet reviews, and customer support channels, among other places. This data analysis is too time consuming and impractical to be done by hand. AI-driven sentiment analysis solves this problem by automatically analyzing and comprehending enormous volumes of textual data through the use of machine learning techniques, natural language processing (NLP), and text analysis. Determining if the feedback is neutral, negative, or favorable is the main objective. Businesses can get rapid insights into customer sentiments, satisfaction levels, and emerging trends by automating this process.



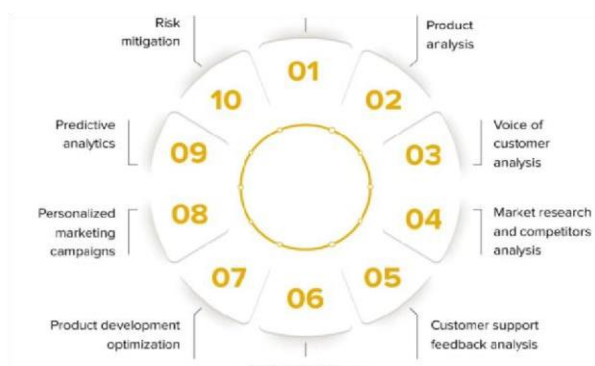
IMPORTANCE OF SENTIMENT ANALYSIS

Retrieving valuable information from customer comments is the main objective of sentiment analysis. Organizations may better assess consumer satisfaction levels, pinpoint are as for development, and build on strengths by identifying whether sentiments are positive, negative, or neutral. Businesses can use this study to make data- driven decisions that enhance their product offerings, customer support, and general reputation as a brand.



ENHANCING ACCURACY WITH SARCASM DETECTION

One of the difficulties in sentiment analysis is recognizing sarcasm, a type of communication in which the intended meaning contradicts the literal meaning of words. Sarcasm detection algorithms seek to identify linguistic cues, context, and patterns that indicate sardonic intent. Sarcasm detection in sentiment analysis techniques is critical because sarcastic remarks might bias results if not correctly identified. This update ensures that sentiment analysis gives reliable insights based on true consumer sentiments rather than incorrect or false data.



INSIGHTS

Sentiment analysis of customer feedback is a useful tool for organizations to better understand and respond to *client attitudes. Companies can use advanced techniques like sarcasm detection to extract subtle insights, improve decision-making processes, and ultimately increase customer happiness and loyalty. As businesses continue to rely on consumer feedback as a key component of their operations, sentiment analysis will become increasingly important in deciphering these insights.*

LITERATURE REVIEW

SENTIMENTAL ANALYSIS

Sentiment analysis has advanced dramatically over the last few decades, from simple rule-based approaches to sophisticated machine learning algorithms. Initially, sentiment analysis was mainly reliant on manual and rule-based methods. These methods used established rules and lexicons to recognize and categorize sentiment in text. Rule-based approaches, while useful for small-scale applications, were hampered by their inability to handle natural language nuances and complexity.

As the area advanced, statistical approaches and machine learning algorithms became more important. These methods used massive datasets and computer capacity to identify patterns and make predictions. Using labeled data, early machine learning algorithms like Naïve Bayes and Support Vector Machines (SVM) were able to outperform rule-based systems.

SARCASM ANALYSIS

Sarcasm identification is intrinsically difficult because of the nuanced character of sarcastic discourse. Sarcasm frequently entails saying something that contradicts what is intended, which can be difficult for algorithms to detect. The challenge is to recognize context, tone, and, in some cases, external knowledge that a human reader would naturally comprehend. Sarcasm, as opposed to pure sentiment expressions, frequently relies on the reader's ability to detect irony, exaggeration, or underlying meanings that contradict the exact words.

Sarcasm detection is complex due to several factors:

1. Context Dependency: Sarcasm frequently relies on context, which may not be explicitly mentioned in the text. Understanding sarcasm frequently necessitates knowledge of past talks, common information, or the context in which the statement is delivered.
2. Subtlety and Ambiguity: Sarcastic remarks can be quite subtle, making them difficult to discern from genuine assertions. The vagueness of such statements hinders the detecting process.
3. Variability in Expression: Sarcasm can be conveyed in a variety of ways, ranging from blatant to subtle. This diversity makes it challenging to develop a one-size-fits all model for detecting sarcasm.

UNDERSTANDING SENTIMENT ANALYSIS

FUTURE APPLICATIONS OF SENTIMENT ANALYSIS

- Predictive Analytics: Sentiment analysis can forecast future trends and consumer behavior by analyzing current textual data.
- Decision Making: Businesses use sentiment analysis to make informed decisions based on customer feedback and social media sentiments.
- Brand Reputation Management: Companies can monitor their brand reputation in real-time and proactively address negative sentiments.

KEY APPLICATIONS AND BENEFITS

- Customer Feedback Analysis: Analyzing reviews and social media comments helps gauge customer satisfaction and pin point areas for improvement.
- Social Media Monitoring: Monitoring platforms provide insights into public opinion, trends, and sentiments related to products and brands.
- Market Research: Conducting sentiment analysis on market trends and consumer preferences informs strategic business decisions.

BENEFITS OF SENTIMENT ANALYSIS

- Enhanced Customer Engagement: Understanding customer sentiments leads to improved engagement and loyalty.
- Real-time Insights: Provides timely information for prompt responses to emerging trends.
- Competitive Advantage: Companies gain an edge by aligning products and services with customer preferences and market sentiment.

BUSINESS PRIORITIES

- Improve customer retention by 10%
- Enhance operational efficiency

ADDED PRIORITIES

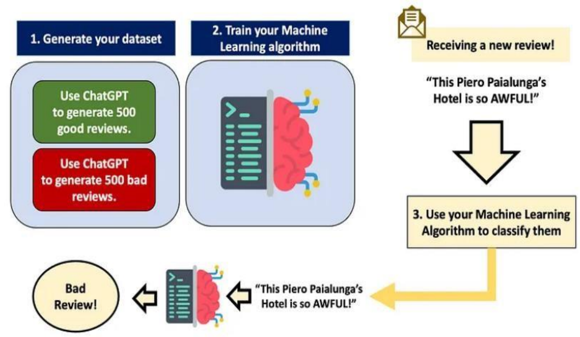
- Strengthen social media presence to boost brand visibility.
- Maintain cost-effective development practices to stay within budget constraints.

FUTURE REFERENCE WORK

Looking ahead, the future of sentiment analysis involves a comprehensive process aimed at understanding and categorizing the emotions expressed in textual data. The initial phase, preprocessing, plays a pivotal role by identifying keywords that encapsulate the core message of the text, followed by tokenization to break down sentences into manageable units or tokens. Lemmatization then standardizes words into their base or root form, facilitating more accurate analysis. Stop-word removal filters out common but in significant words that do not contribute to the overall sentiment.

Moving forward, sentiment analysis employs various approaches, including the rule-based method, which assigns sentiment scores based on predefined lexicons of positive and negative words. This approach simplifies setup and interpretation but requires continual updates to lexicons and may struggle with cultural nuances. Alternatively, leveraging machine learning techniques like neural networks allows for more nuanced emotional detection, where algorithms discern not just positive or negative sentiments but also complex emotions like joy, anger, or regret based on the context and choice of words.

Despite advancements, challenges persist, such as accurately interpreting sarcasm, handling sentences with negation that reverse the sentiment, and addressing multipolarity whereas ingle sentence conveys mixed emotions. These complexities underscore the on going need for sophisticated algorithms capable of understanding subtle nuances in language to provide meaningful insights. By overcoming these challenges, sentiment analysis aims to empower businesses with a deeper understanding of customer opinions, social trends, and market dynamics, enabling informed decision-making and strategic planning to drive organizational success.

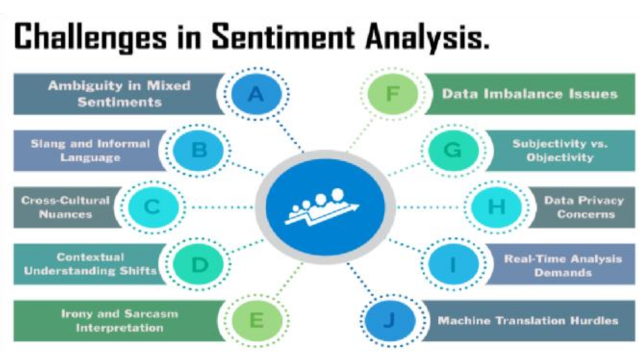


CHALLENGES

Sentiment analysis faces several challenges that impact its effectiveness and reliability in interpreting textual data accurately. One significant challenge is sarcasm detection, where understanding the intended meaning of sarcastic statements proves difficult for algorithms. For instance, phrases like "Oh, great, another meeting" may appear positive at face value but convey negativity.

Negation poses another hurdle as it reverses the sentiment of a sentence, making it challenging for sentiment analysis models to correctly interpret statements such as "I don't dislike this product," which expresses a positive sentiment despite the presence of negation.

Moreover, multipolarity complicates sentiment analysis by presenting sentences that contain multiple sentiments. For example, are view stating" The product features are great, but customer service needs improvement" requires the analysis of each sentiment separately to provide accurate insights.



METHODOLOGY

DATA COLLECTION

1 Gather datasets:

- Dataset1: Sentiment Analysis

Ensure that the dataset includes enough textual data labeled with attitudes (negative, neutral, positive).

- Dataset2: Sarcasm Detection

Check that the dataset contains headlines labeled as sarcastic or non sarcastic.

- Dataset3: Additional Sentiment Analysis.

Use this dataset to validate the sentiment and sarcasm recognition model.

2 Load Dataset:

1. Load Dataset
2. Load dataset2
3. Load Dataset3

DATA PREPROCESSING

3 Text Cleaning:

- Create a function to preprocess text by removing non-word characters with regular expressions and converting it to lowercase.

4 Feature Extraction:

- TF-IDF Vectorization converts textual data into numerical features.

5 Model Development: Sarcasm Detection Model

- Use the Random Forest Classifier algorithm.
- Train the model using the sarcasm detection dataset (Dataset2).
- To evaluate model performance, use appropriate measures like as accuracy, precision, recall, and F1-score.

6 Sentiment Analysis Models:

- Algorithm: Logistic Regression.
- Train the model using the sentiment analysis dataset (Dataset1).
- Evaluate the model's performance using metrics such as accuracy, precision, recall, and F1-score.

7. Sentiment Prediction with Sarcasm Adjustment:

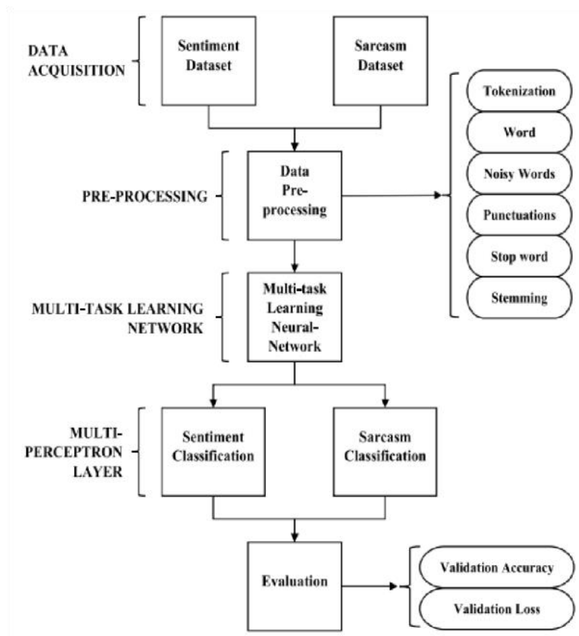
- Create an integrated approach for adjusting sentiment predictions depending on sarcasm detection results.
- Implement logic to change sentiment labels when sarcasm is recognized.

8. Example prediction:

- Provide example sentences and sentiment forecasts before and after sarcasm adjustment.

9PerformanceMetrics:

- Report accuracy and confusion matrix for sentiment analysis model.
- Report accuracy and confusion matrix for detecting sarcasm.
- Validation for Dataset3: Compare projected sentiments to reals entiments in Dataset3 and report accuracy.



DISCUSSION

Model Performance Strengths:

Accuracy:

Both the sentiment analysis and sarcasm detection models perform well on their respective test sets. Given the nature of the datasets, Random Forest Classifier for sarcasm detection and Logistic Regression for sentiment analysis are both viable options.

```
Sarcasm Detection Model Report:
              precision    recall  f1-score   support

           0       0.79      0.82      0.80      2996
           1       0.75      0.72      0.74      2346

    accuracy                           0.77      5342
   macro avg       0.77      0.77      0.77      5342
weighted avg       0.77      0.77      0.77      5342

Sentiment Analysis Model Report:
              precision    recall  f1-score   support

           0       0.68      0.43      0.52       207
           1       0.56      0.80      0.66       286
           2       0.76      0.61      0.68       214

    accuracy                           0.63       707
   macro avg       0.67      0.61      0.62       707
weighted avg       0.66      0.63      0.62       707
```

Feature Extraction:
TheTF-IDF vectorization method successfully captures the value of words in documents, allowing models to discern between different attitudes and detect sarcasm.

Integration Approach:
The adjustment of sentiment predictions based on sarcasm detection is an innovative and useful strategy that improves overall sentiment prediction accuracy by taking into account sarcasm.

Limitations:
Dataset Size and Quality:
The performance of the models is significantly influenced by the quality and amount of the datasets. Limited or biased data can result in poor generalization. Future research could focus on increasing the datasets to include more diverse samples.

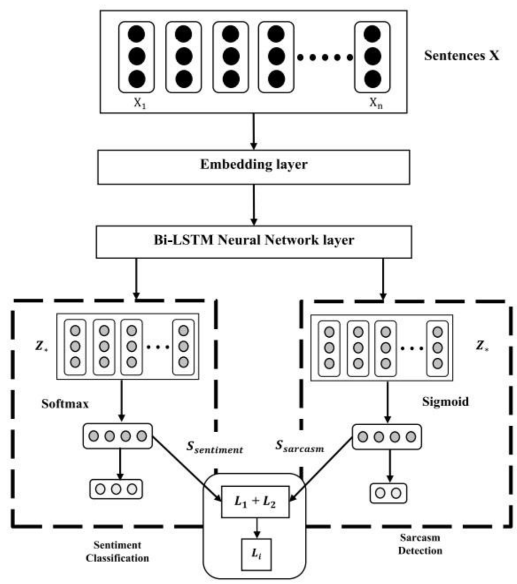
Model complexity:
While the Logistic Regression model for sentiment analysis is successful, its simplicity may limitits capacity to detect more complicated patterns in data. Exploring more sophisticated models, such as deep learning approaches, could improve performance even further.

Sarcasm Detection Challenges:
Sarcasm recognition is still a difficult task due to its delicate and context-dependent character. The Random Forest Classifier, while useful, may not capture all nuances of sarcasm. Incorporating contextual information and advanced NLP approaches may enhance detection rates.

Factors influencing performance:
Feature Extraction Techniques:
The choice of TF-IDF vectorization has a considerable impact on model performance. Alternative approaches, such as word embeddings (Word2Vec, GloVe) and transformer-based embeddings (BERT), could yield more detailed representations of the text.

Hyperparameter Tuning:
Fine-tuning the hyperparameters of the models, notably the Random Forest Classifier, may produce superior results. Grid search and randomized search techniques are suitable for this purpose.

Cross-Validation:
Cross-validation during the training phase can provide a more reliable assessment of model performance and assist in spotting any overfitting or under fitting concerns.



Customer Service:
Sentimental Analysis:
The integrated solution may be used to analyze consumer feedback in real-time, allowing organizations to immediately determine and address negative feelings. This may increase consumer satisfaction and loyalty.

Sarcasm detection:
Satire in customer reviews must be identified for sentiment analysis to be dependable. This method can help customer service representatives perceive sarcastic remarks more accurately, which can improve response quality and minimize misunderstandings.

SOCIAL MEDIA MONITORING

Brand Monitoring:
Businesses can monitor sentiment trends on brand-relevant social media platforms with this tool. Sarcasm is common on social media platforms, so being able to recognize it the reis extremely helpful.

Crisis Management:
The algorithm can help detect possible PR issues early by precisely detecting the sentiment of social media remarks, including sarcastic cones.

FUTUREWORK

Model improvements:
Future studies could look into advanced models like deep learning, particularly LSTM or transformer-based models, to improve sentiment and sarcasm recognition capabilities.

Contextual Understanding:
Context-aware models, which take into account surrounding text or conversation history, can enhance sentiment and sarcasm recognition accuracy.

Multilingual Support:
Expanding the system to handle several languages can increase its application and usefulness in global settings.

Real-Time Processing:
Optimizing the system for real-time processing can make it appropriate for applications that require fast feedback, such as live chat help or social media monitoring.

RESULT

Sarcasm Detection Model:
An accuracy of 0.77 is quite good, especially given the challenges of detecting sarcasm, which often involves understanding nuanced language and context.
The precision and recall for both classes are fairly balanced, indicating that the model performs consistently across both sarcastic and non-sarcastic instances.
The F1scores of 0.80 and 0.74 suggest that the model is performing well, with a slight drop in the sarcastic class, which is expected given the complexity of sarcasm.

Sentiment Analysis Model:
An accuracy of 0.63 indicates a moderate level of performance.
The precision and recall vary significantly between classes, with neutral sentiment being detected more reliably (recall of 0.80) compared to negative sentiment (recall of 0.43).
The lower recall for negative sentiment suggests that many negative instances are being misclassified, which could be problematic if accurate detection of negative feedback is critical.
The F1-scores of 0.52 (negative), 0.66 (neutral), and 0.68 (positive) show that the model has room for improvement, particularly in distinguishing negative sentiment.

CONCLUSION

The code creates a strong foundation for sentiment analysis and sarcasm detection, utilizing machine learning techniques like Random Forest Classifier for sarcasm detection and Logistic Regression for sentiment analysis. This arrangement allows for a variety of applications, such as analyzing consumer comments and tracking sentiment across social media sites. The code enables accurate and nuanced interpretation of textual data by preprocessing text, vectorizing characteristics, and forecasting sentiments, hence aiding informed decision-making and strategic in sights across multiple domains and also focuses on the below things

Libraries Imported: For data handling, text processing, machine learning, and sentiment analysis, the code makes use of numerous powerful libraries such as pandas, re, sklearn, and text blob.

Data Preprocessing: Text from various datasets is preprocessed to remove non-word characters and convert them to lower case. This standardization is required for proper text analysis.

Dataset Loading: Three different data sets are loaded from CSV, JSON, and Excel files to demonstrate the code's ability to handle a variety of data types.

Vectorization: Tfidf Vectorizer transforms text data into numerical features, which is an important step for machine learning models when processing text data.

Model Training:
A Random Forest Classifier is taught to identify sarcasm. A Logistic Regression model is trained to perform sentiment analysis.

Sentiment Prediction: A function called predict sentiment is written to predict the sentiment of fresh text inputs. This method uses sarcasm detection to change the sentiment forecast, showcasing a more subtle approach to text analysis.

Example Prediction: The sentiment prediction technique is demonstrated with an example statement, including how to handle sarcasm.

Application to New Dataset: The trained models are applied to a new dataset (Dataset3), predicting sentiments for each text entry and comparing the predictions to real sentiments if available.

Result Storage: The final findings, including anticipated attitudes, are saved in a new Excel file, ensuring that the output is documented for future study or reporting.

Accuracy Calculation: If real feelings are available, the algorithm calculates and outputs the forecast accuracy, which serves as a measure of model performance.