Product Review and Rating Classification Using Sentiment Analysis

Group Name

Neural Nex

Group Members

Ritesh Sengar

Shravan Honade

Nikhil Patil

Dhananjay Ghate

Agenda

- Introduction
- Understanding Sentiment Analysis
- Machine Learning Models for Sentiment Analysis
- Parameters for Product Review and Rating Classification
- About Data
- Data Collection and Preprocessing
- Training and Evaluation
- Classification of Product Reviews
- Challenges in Sentiment Analysis
- Applications and Benefits
- Conclusion

Introduction

Product reviews are valuable feedback from customers, and rating classification involves categorizing these reviews based on sentiment or performance. Sentiment analysis helps to determine the sentiment expressed in the text, such as positive or negative. Machine learning models are algorithms that can learn from data to make predictions or classifications. Let's dive into the details of this interesting field.

General Terms

- **Product Review:** Reviews provided by customers about product or service.

- **Rating Classification:** The process of categorizing product reviews into groups based on sentiment or performance.

- **Sentiment Analysis:** The technique to determine the sentiment expressed in text, such as positive, negative, or neutral.

- **Machine Learning Model:** Algorithms that can learn from data to make predictions or classifications.

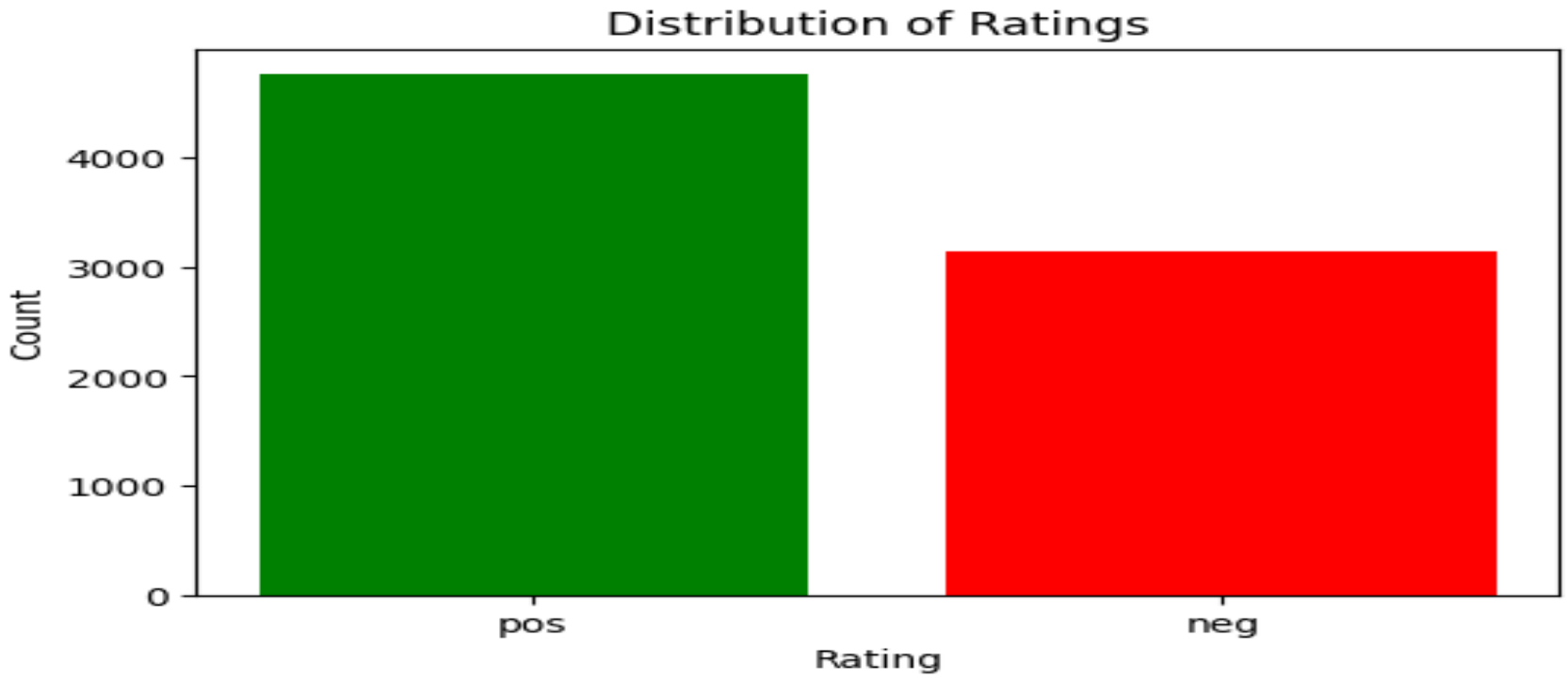
Understanding Sentiment Analysis

- **Definition:** Sentiment analysis is the process of determining the sentiment expressed in text, such as positive, negative, or neutral.
- **Process:** The process of sentiment analysis involves text text preprocessing, feature extraction, model training, and sentiment classification.
- **Importance:** Sentiment analysis is important for understanding customer opinions, improving product/services, and making data driven decisions.

Parameters for Product Review and Rating Classification

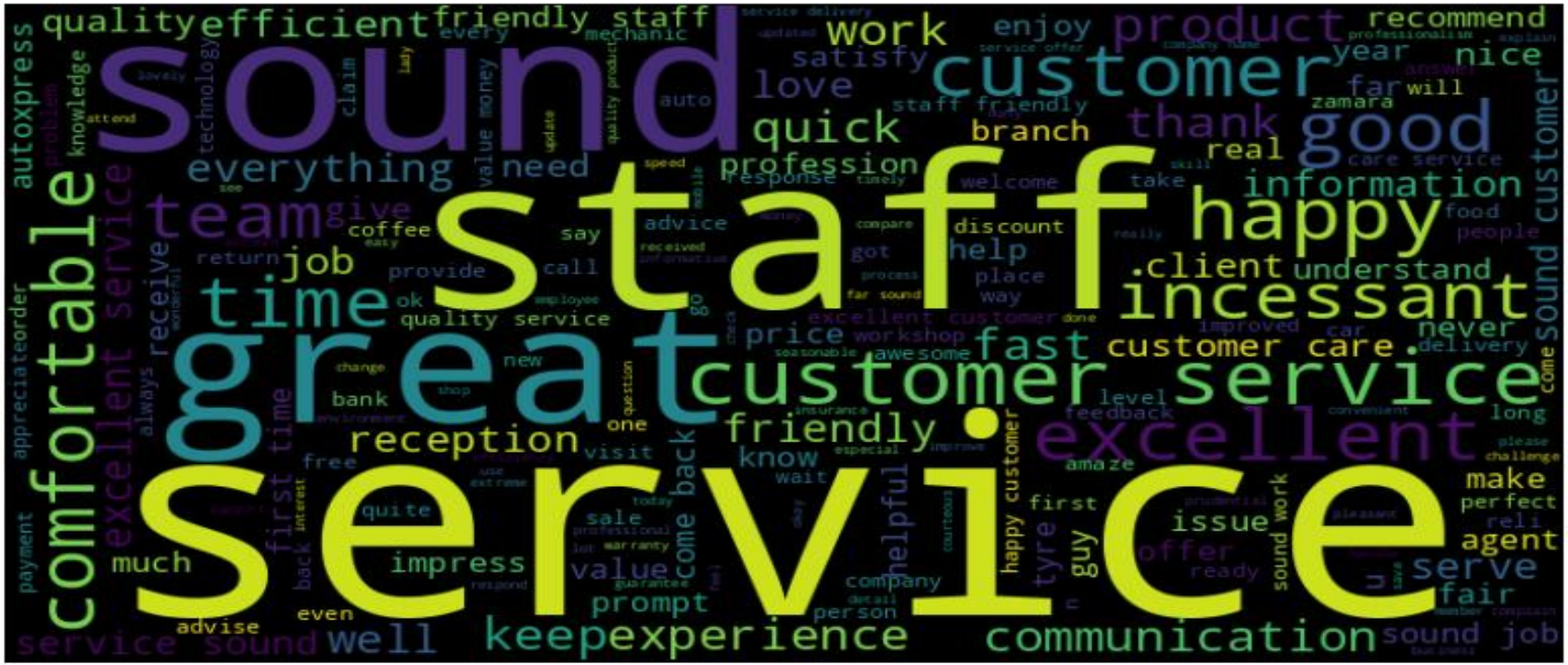
- **Quality of Product**: Evaluated based on the features, performance, durability, and user experience of the product.
- **Quality of Communication**: Assesses the clarity, responsiveness, and effectiveness of customer communication channels.
- **Quality of Staff**: Refers to the knowledge, friendliness, and helpfulness of the staff interacting with customers.
- **Value for Money**: Measures whether the product or service provides good value in relation to its price.

About Data


Distribution of Ratings
Count
4000
3000
2000
1000
0
pos
neg
Rating

About Data

Word Cloud for Positive Sentiment



Data Collection and Preprocessing

- **Crawling Product Review Data**: Collecting product review data from various sources, such as websites or social media platforms.
- **Text Cleaning**: Removing noise, unwanted characters, and irrelevant information from the text data.
- **Tokenization**: Splitting the text into individual tokens or words to facilitate further Analysis.
- **Profanity:** Profanity filtering is a process used to automatically detect and filter out offensive, vulgar, or inappropriate language, known as profanity, from text or speech data.

Data Collection and Preprocessing

- **Stemming**: Stemming is a text normalization process in natural language processing (NLP) and information retrieval. Its primary purpose is to reduce words to their root or base form, called the "stem."

- **Missing values**: Punctuation and missing values handling.

- **TF/IDF Vectorization**: "TF-IDF vectorization is a numerical representation technique used in natural language processing to convert a collection of text documents into numerical vectors, where each vector represents a document's importance in relation to specific terms, combining Term Frequency (TF) and Inverse Document Frequency (IDF) information."

Vectorization

1. Julie loves John more than Linda loves John
2. Jane loves John more than Julie loves John

```
John   2    2
Jane   0    1
Julie  1    1
Linda  1    0
likes  0    1
loves  2    1
more   1    1
than   1    1
```

the two vectors are,

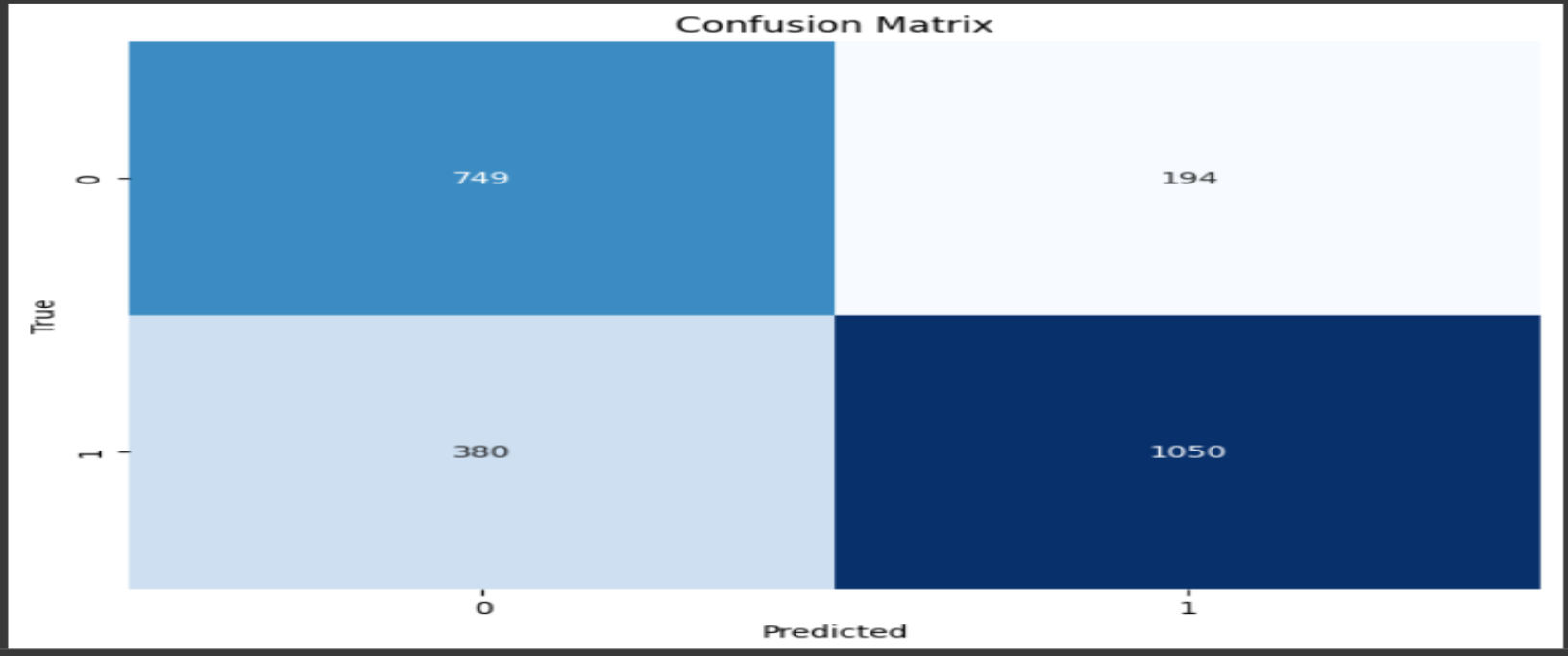
```
Item 1: [2, 0, 1, 1, 0, 2, 1, 1]
Item 2: [2, 1, 1, 0, 1, 1, 1, 1]
```

Training and Evaluation

```
]  Accuracy: 75.81%

   Confusion Matrix:
   [[ 749  194]
    [ 380 1050]]
```

****



Classification of Product Reviews

```
The way of talking of staff is not efficient
Performing Well : []
Needs Improvements : ['Quality of staff']
Sentiment of the Sentence :  Negative

**************************************************
The quality of product is good but communication reports are bad
Performing Well : ['Value for money', 'Quality of product']
Needs Improvements : ['Quality of communication']
Sentiment of the Sentence :  Positive

**************************************************
Support staff is understanding and co-operative
Performing Well : ['Quality of staff']
Needs Improvements : []
Sentiment of the Sentence :  Positive

**************************************************
```

Category Tagging

Category tagging, also known as content categorization or topic classification, is a natural language processing (NLP) technique used to assign predefined categories or tags to a piece of text or content.

```
**************************************************
The quality of product is good but communication reports are bad
Performing Well : ['Value for money', 'Quality of product']
Needs Improvements : ['Quality of communication']
Sentiment of the Sentence :  Positiv
```

Challenges in Sentiment Analysis

Sarcasm: Sarcasm often involves the use of words or phrases that have a different meaning from their literal interpretation.

Irony: Irony refers to the use of words or phrases to express the opposite of their literal meaning.

Contextual Understanding: Sentiment analysis models often struggle With understanding the context of text. The meaning of words and phrases can vary depending on the surrounding context. For example, the sentiment of the word ’bad' can change if it is used in the context of 'not bad' or 'very bad’. Improving contextual understanding is crucial for accurate sentiment classification.

Applications and Benefits

APPLICATIONS:

Customer Feedback Analysis: Customer Feedback Analysis, where sentiment analysis can be used to analyze customer reviews and feedback to gain insights into customer satisfaction and identify areas for improvement.

Brand Reputation Management: Brand Reputation Management, where sentiment analysis can help monitor and analyze online conversations to understand public sentiment towards a brand or product.

BENEFITS:

By leveraging sentiment analysis, businesses can make data-driven decisions, improve customer experience, and enhance their brand reputation.