Sentiment Analysis using Machine Learning

Logo

Description automatically generated

CSCI417-Machine Intelligence

Final Report

Salma Khaled AbdelSamiee (19100068)

Andrew Effat Lotfy (19105700)

Table of Contents

**Table of Contents1**

Abstract2

Introduction3

Methodology4

What is Sentiment Analysis?4

Types of Sentiment Analysis4

Fine-Grained4

Emotion Detection4

Aspect Based5

Multi-lingual5

Importance of Sentiment Analysis?5

Implementation6

**Results8**

Conclusion12

References13

1. **Abstract**

Sentiment analysis is growing more and more popular among businesses that seek to determine the emotions of their clients because of technological breakthroughs in areas like deep learning. Businesses now classify words into positive, negative, and neutral categories using natural language processing, statistical analysis, and text analysis to determine the sentiment. The top businesses recognize how crucial it is to comprehend the feelings of their clients, including what they are saying, what they mean, and how they are saying it. When people mention your brand in comments, reviews, tweets, or other social media platforms, you may utilize sentiment analysis to determine how people feel about it. We have created a comprehensive guide to help you understand "what is sentiment analysis?" its tools, and its various categories and use cases since sentiment analysis is the field of utilizing software to interpret emotions. [1]

1. **Introduction**

Sentiment analysis has developed into a potent tool for tracking and comprehending online dialogues as consumers share their opinions and thoughts about the brand more freely than ever before. You may discover what makes a customer happy or unhappy by automatically analyzing feedback and reviews from surveys or social media conversations. Additionally, you may utilize this information to customize your goods and services to match the needs of your clients and strengthen your brand. The effectiveness of sentiment analysis algorithms has risen because of recent developments in machine learning and deep learning. Innovative machine learning and advanced artificial intelligence techniques can be used to conduct research and produce analysis. For instance, sentiment analysis can assist you in automatically analyzing more than 5000 reviews of your business to determine whether your customers are pleased with your pricing strategies and level of customer support, or not. As a result, you may assert that there are countless applications for emotion. [1]

1. **Methodology**
2. **What is Sentiment Analysis?**

We have mentioned about the benefits of sentiment analysis, but it was not mentioned what is actually sentiment analysis.

Analysis of the customer's positive or negative attitude in text is known as sentiment analysis. Businesses can monitor online chats to use contextual analysis of identifying information to better understand the social attitude of their customers.[1]

1. **Types of Sentiment Analysis**
2. **Fire-grained:**

You can divide your polarity categories into the following categories if your business demands polarity precisions:

* Positive
* Neutral
* Negative
* Very Negative

Use the 5-star ratings as a customer evaluation for polarity analysis, where extremely positive is a five-star rating and very negative is a one-star rating. [1]

1. **Emotion Detection**

This kind of sentiment analysis aids in the identification of client emotions such as joy, disappointment, rage, grief, etc. To determine the customer's feelings in this situation, you can employ sophisticated machine learning algorithms or sentiment lexicons. The fact that people frequently communicate their feelings in a variety of ways is one drawback of employing sentiment lexicons. Therefore, it could be difficult to fully perceive human emotion while utilizing it. [1]

1. **Aspect Based**

Consider that you are employing fine-grained analysis to examine customer sentiment. You want to figure out which specific feature or functionalities are receiving positive or negative feedback. Aspect-based sentiment analysis is used in this situation. For instance, an aspect-based classifier will determine that the review "The camera quality of this phone is growing worse with time" reflects a dissatisfaction with the phone's photography feature. [1]

1. **Multi-lingual**

Multilingual sentiment analysis is more challenging than other varieties since it needs substantial preprocessing and web resources (i.e., sentiment lexicons). Businesses value customer feedback regardless of the customer's origin or linguistic proficiency. Therefore, multilingual sentiment analysis helps to determine client sentiment regardless of geography or linguistic variances. [1]

1. **Importance of Sentiment Analysis**

The most important benefit of sentiment analysis is that it makes it possible for you to comprehend how your clients feel about your brand. By automatically assessing the thoughts and sentiments of your customers through social media chats, reviews, polls, and more, you can enhance your goods and services and make better decisions. [1]

According to our survey, 23 people responded, and they all agree that sentiment analyzer is beneficial, and also, they suggested for whom is the analyzer beneficial for.

Chart

Description automatically generated

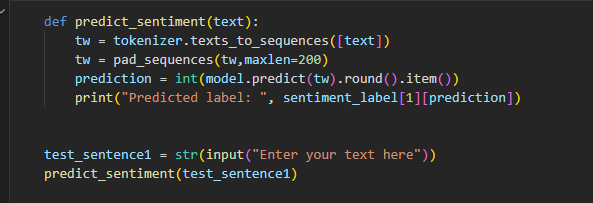
1. **Implementation Steps [2]**

This is the implementation steps of binary classification fine grained sentiment analysis:

1. Data Preprocessing
2. First, we preprocess the data set as we are dealing with text
3. We extract only the columns with text reviews, and their sentiments.
4. We are going to analyze positive or negative texts only, so the rows that contain neutral sentiment will be dropped out.
5. The labels (positive. negative) are text, and machine learning algorithms only understand numeric values, so factorize method is used to convert these texts to numeric values, returning an array of numeric values with their matching indexes (0 is positive, 1 is negative).
6. We extract the text column from the dataset and store them in a variable, to tokenize the texts. Tokenization means breaking down each text to small parts called tokens. Then fit\_on\_texts method is used to match the words with their assigned number. This matching is stored in a dictionary, so then these words are replaced with their matching numbers using texts\_to\_sequence method. Finally, we use pad\_ sequences method to make all sentences have the same length.
7. Text Classifier
8. In the machine learning model for the sentiment analysis project, we use LSTM layers. Three layers make up the architecture of our model: an embedding layer, an LSTM layer, and a dense layer. To avoid overfitting, we inserted the Dropout mechanism in-between the LSTM layers. LSTM stands for Long Short-Term Memory Networks. Recurrent Neural Networks are a version of it. With sequential data, such as text and audio, recurrent neural networks are typically used. Typically, when calculating an embedding matrix, hidden states—which refer to the computations made for each word—are stored. RNNs are unable to store all these calculations in their memory if, for example, a word is referenced 100 words into a text. RNNs are therefore unable to learn these long-term dependencies.
9. One method of regularization is dropout. It's employed to prevent overfitting. In the dropout method, we erratically remove certain neurons. A number between 0 and 1 that specifies the likelihood that the neurons will be dropped is the argument that the layer accepts. This develops a solid model while preventing overfitting.
10. Model Training
11. Model Execution
12. **Results/Discussion**

Model 1:

The user enters any text of his own, the function first tokenizes the text to break it down to tokens, padding these tokens to have the same length. Finally we predict the sentiment of the input text , and the put is then displayed.



* 1. Positive Text:

The input text here was ”I am over the moon”. This is an idiom which expresses how someone is very happy. The output should be positive.

Graphical user interface, application, Teams

Description automatically generatedText

Description automatically generated

* 1. Negative Text:

The input text here was” I feel like having a lump in my throat”. This is an idiom which expresses how someone is very sad and heart broken. The output should be negative.

Graphical user interface, application, Teams

Description automatically generated

A picture containing text

Description automatically generated

**Model 2:**

**A screenshot of a computer

Description automatically generated**

* Logistic Regression Mode:

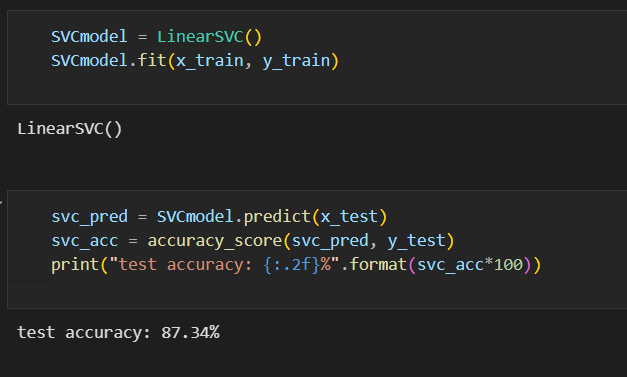
Text

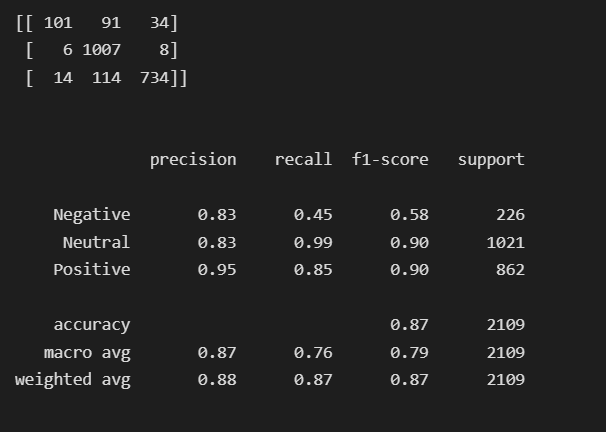
Description automatically generated

Calendar

Description automatically generated

* Linear SVC Mode:





1. **Conclusion**

Due to the development of technology, the era of gaining insightful data from surveys and social media has come to an end. As a result, it is now important for your company to be aware of how your clients are feeling. To maximize the potential of data and gain the most insightful understanding, businesses are utilizing sophisticated classifiers like sentiment analysis and contextual semantic search. Utilizing technologies for natural language processing, you may create business strategies, meet, or exceed customer expectations, generate leads, create marketing campaigns, and uncover new growth opportunities. [1]

1. **References**

[1] “Introduction to Sentiment Analysis: Concept, Working, and Application,” *Maruti Techlabs*, Oct. 08, 2021. https://marutitech.com/introduction-to-sentiment-analysis/

[2] “Sentiment Analysis using Python [with source code],” TechVidvan, Jun. 14, 2021. https://techvidvan.com/tutorials/python-sentiment-analysis/