

**Chandigarh School of Business Jhanjeri,Mohali-140307**

**Department of Computer Applications**

Major Project on

**SENTIMENT ANALYSIS**

**BACHELOR OF Computer Application**



**SUBMITTEDBY:**

Name: Harsh Nayyar

Rollno:2133078

Name: Jaswinder

Rollno:2133089

(MAY2024)

**Under The Guidance Of**

Name : Naveen Bansal

Department of ComputerApplication

Chandigarh School Of Business,Jhanjeri

Mohali - 140307



**Chandigarh School of Business Jhanjeri,Mohali-140307**

**Department of Computer Applications**

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **SNo** | **Contents** | **Page No** |
| 1. | Introduction | 1 |
| 2. | Abstract System | 2 |
| 3. | Existing System | 3 |
| 4. | Proposed System | 4 |
| 5. | Objectives | 5 |
| 6. | Hardware and Software Requirements | 6 |
| 7. | Conclusion | 7 |
| 8. | References | 8 |

INTRODUCTION

Data is being produced at an astounding rate and volume in the field of the internet and other digital services nowadays. Researchers, engineers, and data analysts often work with tabular or statistical data. There may be categorical or numerical data in each of these tabular data columns. Various data formats, including text, picture, video, and audio, are present in data that is generated. Analysis of unstructured data is produced by online behaviour such as publications, web content, blog entries, and social media platforms. To effectively build their business, corporations and businesses ONE must examine textual data to comprehend consumer behaviours, opinions, and comments. Text analytics is developing at a higher pace in order to deal with massive text information.

Sentiment Analysis: what is it?

The method of determining how well a chunk of content is good, negative, or neutral is known as sentiment analysis. Sentiment analysis is just the contextually extraction of words that reveals the social attitude of a brand and aids businesses in determining if the products they are producing will find a market. Sentiment analysis's objective is to examine public sentiment in a manner that will support corporate growth.

What is the Process of Sentiment Analysis?

The Automatic Method is the first strategy. This tactic makes use of machine learning. Once the datasets have been analysed, predictive analysis is then carried out. The next step is word extraction from either the text. Text analysis can be done using a variety of techniques, including Naive Bayes classifier, Regression Analysis, Support Vector, and machine learning algorithms.

Rule-based strategy is the second strategy. The lexicon method, rule-based tokenization, and parsing are all applied in this case. The method counts the number of positive and negative phrases in the sample. The emotion is positive if there are more upbeat than downbeat messages; otherwise, it is the opposite.

The most accurate method for sentiment analysis is Hybrid Approach. This method combines the rule-based and automated procedures described above. The advantage is that, in comparison to other major procedures, accuracy will be great.

ABSTRACT SYSTEM

Sentiment analysis is a method for figuring out the emotional undertone or sentiment expressed in textual information. Customer feedback analysis, brand reputation management, market research, and social media monitoring are just a few of the many fields in which it has applications. This abstract system describes the main elements and steps in sentiment analysis.

Components:

1.Data collection: Compile textual information from a variety of sources, including internet forums, social media sites, customer reviews, and surveys. Sentiment analysis will use this data as its input.

2.Data cleaning and preprocessing: Remove noise and unrelated data from the collected data. To establish a clean and uniform dataset, procedures such text normalisation, stop word removal, addressing spelling problems, and tokenization are carried out.

3.Create numerical features from the preprocessed text that machine learning or deep learning models can use to process the content. Word embeddings like Word2Vec or GloVe, TF-IDF, or bag-of-words approaches are frequently used for feature extraction.

4.Machine learning or deep learning methods can be used to train a sentiment analysis model. To categorise text into categories like positive, negative, or neutral emotion, the model learns from instances that have been labelled. Naive Bayes, Support Vector Machines, and neural network topologies like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) are examples of popular algorithms.

5.Model assessment: Utilise performance evaluation criteria including accuracy, precision, recall, and F1 score to rate the trained sentiment analysis model's effectiveness. This stage guarantees that the model can accurately categorise feelings in unobserved data and provide information on its advantages and disadvantages.

6.Deployment and Integration: Make the sentiment analysis model available for real-time sentiment analysis by deploying it in a production environment. To process and analyse fresh textual data, this can entail creating APIs, incorporating the model into current systems, or creating web-based apps.

7.Continuous Improvement: Track user feedback and continuously assess the effectiveness of the implemented model. To react to changing language patterns, increase accuracy, and meet domain-specific sentiment analysis requirements, gather fresh labelled data and retrain the model periodically.

EXISTING SYSTEM

The current system of sentiment analysis includes a number of methods and procedures used to evaluate and categorise sentiment in textual data. Here are a few typical techniques for sentiment analysis:

1.Rule-Based Approaches: To detect sentiment in text, rule-based algorithms use predetermined sets of linguistic rules and patterns. These guidelines may be written by hand or obtained from lexical or lexical resources. In rule-based techniques, the presence of positive or negative words and their context are frequently used to assign sentiment scores or labels.

2.Techniques for Machine Learning: Sentiment analysis frequently employs machine learning methods. When trained on labelled data, supervised learning algorithms like Naive Bayes, Support Vector Machines (SVM), and Decision Trees can discover patterns and connections between text attributes and sentiment labels.These models can then classify unseen text into sentiment categories.

3.Lexicon-Based Approaches: Lexicon-based approaches make use of dictionaries or lexicons that have words or sentences that have the polarity of the attitude (positive, negative, or neutral) noted. Based on the vocabulary's presence and polarity, the sentiment of a given text is determined. Lexicon-based sentiment analysis often use methods like WordNet, SentiWordNet, or Vader.

4.Hybrid Approaches: To increase the accuracy of sentiment analysis, hybrid approaches integrate several techniques, such as rule-based and machine learning. For general sentiment categorization, a system might utilise rule-based techniques, and for fine-grained sentiment analysis or aspect-based sentiment analysis, it might use machine learning algorithms.

5.Deep Learning Methods: Deep learning models, particularly neural networks, have showed potential in sentiment analysis. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and their variations, such as Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU), have been used to capture complex linguistic patterns and dependencies in text, increasing sentiment classification performance.

6.Aspect-Based Sentiment Analysis: Aspect-based sentiment analysis goes beyond general sentiment categorization and seeks to determine sentiments towards certain aspects or features stated in the text. In order to gain more detailed insights into client attitudes, this strategy entails identifying relevant characteristics or entities and linking sentiment with each component.

7.Continuous Improvement: Track user feedback and continuously assess the effectiveness of the implemented model. To react to changing language patterns, increase accuracy, and meet domain-specific sentiment analysis requirements, gather fresh labelled data and retrain the model periodically.

PROPOSED SYSTEM

Sentiment analysis, which identifies the emotional undertone of a string of words, is frequently employed to ascertain the sentiment of a specific text or document. A suggested system for sentiment analysis is as follows:

1.Data gathering: Compile a dataset containing labelled instances, each of which is assigned a sentiment category (such as positive, negative, or neutral). To ensure the generalisation of the model, the dataset should cover a variety of themes and situations.

2.Clean up and preprocess the data that has been acquired to get rid of extraneous information, such as special characters, punctuation, and irrelevant data. To build a normalised text corpus, carry out actions like tokenization, stemming, and deleting stop words.

3.Feature Extraction: Transform the preprocessed text into numerical features that the sentiment analysis model can utilise as input. Bag-of-words, TF-IDF (Term Frequency-Inverse Document Frequency), or word embeddings like Word2Vec or GloVe are examples of common approaches.

4.Select the right machine learning or deep learning model for your sentiment analysis needs. Naive Bayes, Support Vector Machines (SVM), Random Forests, and deep learning models like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), are some of the most well-liked alternatives.

5.Split the preprocessed dataset into training and testing sets for the model. Using the training set, run the chosen model, and adjust hyperparameters as needed. Based on the input features, the model should develop the ability to forecast sentiment labels.

6.Model Evaluation: Use the testing set to gauge the effectiveness of the trained model. Accuracy, precision, recall, and F1 score are frequently used evaluation measures for sentiment analysis. Make sure to evaluate each sentiment category's performance for the model separately.

7.Deploy the model in a production setting where it can take in new input texts and forecast sentiment after it has been trained and validated. This can be done through APIs, web apps, or integration with other systems.

8.Continuous Improvement: Monitor the performance of the deployed model and collect feedback from users. Continuously update and retrain the model using new data to improve its accuracy and adapt to changing sentiment patterns.

OBJECTIVES

The OBJECTIVES of SENTIMENT ANALYSIS include:

1.Classification of Sentiment: The main goal of sentiment analysis is to categorise the emotional tone or sentiment of a given text as either positive, negative, or neutral. This aids in analysing the general sentiment expressed by people towards a certain product, service, event, or topic.

2.Sentiment analysis gives businesses the ability to examine client input from a variety of sources, including social media, online reviews, client surveys, and support tickets. By evaluating the sentiment of these feedbacks, firms can detect patterns, sentiments, and trends to get insights into customer satisfaction levels, identify areas for development, and make informed business decisions.

3.Brand Monitoring and Reputation Management: Sentiment analysis helps monitor the online reputation of a brand by assessing sentiment expressed in discussions, reviews, and social media dialogues. By analysing sentiment over time, firms can examine the impact of marketing campaigns, product launches, or public relations activities on their brand perception and take proactive measures to manage their reputation.

4.Sentiment analysis offers useful insights into consumer preferences, attitudes, and sentiments towards particular goods, services, or brands, according to market research and competitive analysis. By analyzing sentiment data, organizations can identify emerging trends, conduct market research, assess customer sentiment towards competitors, and uncover opportunities for product innovation or market positioning.

5.Social Listening and Customer Engagement: Sentiment analysis allows firms to actively listen to social media conversations and communicate with customers in real-time. Companies may discover consumer issues and address them quickly, deal with unfavourable sentiment, and capitalise on positive sentiment to build brand loyalty and customer relationships.

6.Risk Assessment and Crisis Management: Sentiment analysis can be effective in spotting possible hazards and crises before they worsen. By monitoring sentiment in social media, news stories, or public forums, organizations can notice early signals of bad sentiment, identify concerns or controversies, and take preemptive efforts to limit risks and manage crises successfully.

7.Opinion Mining and Public Opinion Analysis: Sentiment analysis can be used to gauge public opinion on numerous topics, such as political campaigns, societal issues, or policy reforms. By studying sentiment expressed in public conversation, sentiment analysis can provide insights into public sentiment, track shifts in opinion, and help organizations or policymakers understand popular sentiment towards certain topics.

HARDWARE AND SOFTWARE REQUIREMENTS

Depending on the scope of the analysis, the size of the dataset, and the method (e.g., machine learning, deep learning), the hardware and software requirements for sentiment analysis can change. Here are some broad ideas to keep in mind:

Hardware specifications

CPU: For large-scale sentiment analysis activities in particular, a multi-core CPU (e.g., Intel Core i5 or higher) is advised to manage the computational load effectively.

Speicher (RAM): To accommodate the dataset and the computational needs of the selected models, enough RAM is needed. A minimum of 8GB RAM is normally advised, but larger datasets or complicated models may require more.

Storage: The dataset, preprocessed data, trained models, and any intermediate files produced during the study must all be kept in an adequate amount of space. The storage volume must be based on the size of the dataset and other storage requirements.

Software prerequisites

You'll need a programming language to put the sentiment analysis system into action. Java is a well-liked option because of its extensive ecosystem of libraries and frameworks for machine learning and natural language processing, quality.

Development Environment: Jupyter Notebook, or Visual Studio Code are examples of integrated development environments (IDEs) that can make it easier to develop, debug, and collaborate on code.

Libraries for Sentiment Analysis: Depending on your strategy, you could require particular libraries for sentiment analysis. For instance, NLTK offers a range of resources and tools for natural language processing, such as sentiment analysis. As an alternative, you can create and train sentiment analysis models using deep learning or machine learning frameworks like stanfordnlp, lingpipe,wekaetc.

Tools for Data Preprocessing: Text cleaning, tokenization, and feature extraction are a few examples of data preprocessing techniques that are frequently used in sentiment analysis. For these  meaning cloud,stanford corenlp will be a best suitable tool.

Frameworks for machine learning and deep learning: If you're utilising these kind of models, you'll require some programmes like stanfordnlp, lingpipe,weka etc. These frameworks include pre-built model structures, algorithms, and tools for deployment, training, and evaluation.

other Libraries: If you intend to design a web-based sentiment analysis system, you may require other libraries for tasks like data visualisation (e.g., Matplotlib, Seaborn) or web development (e.g., Flask, Django). This will depend on your individual requirements.

CONCLUSION

The sentiment or emotional tone portrayed in textual data can be understood and analysed using the useful approach known as sentiment analysis. It has numerous uses in various businesses, including reputation management, market research, brand monitoring, and customer feedback analysis.

Organisations can use a sentiment analysis project to extract useful information from vast amounts of textual data, empowering them to make deft decisions, boost customer satisfaction, manage their brand reputation, and spot market trends. Understanding customer sentiment towards particular goods, services, or brands, as a result of sentiment research, enables businesses to adjust their strategies and product offers.

REFERENCES

<https://www.javatpoint.com/sentiment-analysis-using-nltk>

<https://www.geeksforgeeks.org/what-is-sentiment-analysis/>

<https://www.tutorialspoint.com/natural_language_processing/natural_language_processing_quick_guide.htm>