



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

AIKNAP

Arshadali Sayyed¹, Komal Pal², Nikhil Soni³, Prashant Singh⁴, Nilima Kulkarni⁵

Department of Computer Science & Engineering, MIT ADT University, Maharashtra, India

ABSTRACT: Sentiment analysis, the process of extracting emotional signals from textual data, has garnered significant attention owing to its diverse applications in understanding sentiment expressed in digital content. This paper delves into the exploration of sentiment analysis techniques utilizing deep learning libraries such as Natural Language Processing (NLP), Natural Language Toolkit (NLTK), and Long Short-Term Memory (LSTM) networks for website development. The primary objective of this project is to amalgamate computer vision and NLP methodologies to craft an intuitive online application tailored for real-time sentiment analysis of social media data. By categorizing data into positive, neutral, and negative sentiments, the application facilitates insightful analysis, providing valuable insights into user sentiment dynamics. Additionally, the project aims to extend this functionality by developing a website extension dedicated to fostering awareness about mental health issues through the organization of input data and detection of emotional patterns. Through this research endeavor, we seek to contribute to the advancement of sentiment analysis methodologies and their practical applications in website development and mental health advocacy.

KEYWORDS: Sentiment Analysis, Natural Language Processing (NLP), Natural Language Toolkit (NLTK), Long Short-Term memory (LSTM), Deep learning (DL).

I. INTRODUCTION

In the rapidly evolving landscape of business, navigating the challenges posed by the abundance of unstructured text data has become imperative in this digital age. Businesses strive for excellence, aiming to optimize profits and ensure customer satisfaction. Understanding the insights hidden within data has emerged as a crucial task for businesses seeking to thrive in this competitive environment. When launching new products, businesses rely heavily on customer reviews to gauge product performance and identify areas for improvement. However, manually analyzing reviews across various shopping and social media platforms can be a daunting and time-consuming task.

To address this challenge and drive business success while enhancing customer satisfaction, a robust sentiment analysis solution is indispensable. This paper proposes a methodology leveraging the Natural Language Toolkit (NLTK) to develop an integrated model that facilitates text classification, summarization, and interpretation. The proposed model seamlessly integrates into websites, providing users with a user-friendly interface to analyze text and product reviews efficiently.

By employing sentiment analysis, businesses can gain valuable insights into customer sentiments and preferences. Positive sentiments, indicated by words such as "happy" or "joyful," denote favorable reviews, while negative sentiments, characterized by words like "bad" or "cheap," signal areas for improvement. Neutral sentiments, devoid of explicit positive or negative cues, suggest a mixed or neutral sentiment.

Moreover, the proposed sentiment analysis solution extends its functionality beyond analyzing individual reviews to inspecting entire websites. By parsing the content of websites, the system can provide sentiment analysis for the website as a whole, offering users a comprehensive understanding of the overall sentiment surrounding a brand or product. This holistic approach not only saves time by automating the analysis process but also enhances the user experience by providing valuable insights at a glance. With this feature, users can quickly assess the sentiment surrounding a website, enabling them to make informed decisions and enhancing their overall browsing experience.

This proposed methodology empowers customers to make informed purchasing decisions while enabling businesses to refine and enhance their products based on customer feedback. By leveraging sentiment analysis, businesses can adapt to customer preferences, drive product innovation, and ultimately, propel business growth.

Table -1: Literature Survey

Sr. no.	Year	Title	Algorithms/Technique	Performance	Remark
1	2023	An Improved Model for Sentiment Analysis on Luxury Hotel Review	NLP, Deep Learning	To enhance customer satisfaction, the paper proposes employing sentiment analysis to evaluate reviews of high-end hotels, focusing specifically on location, staff training, and room cleanliness.	This paper underscores the significance of scrutinizing extensive customer data by offering an integrated examination of luxury hotel reviews sourced from Booking.com, alongside the introduction of a hybrid model for sentiment analysis tailored to hotel reviews.
2	2023	A BERT Framework to Sentiment Analysis of Tweets	CNN, LSTM, RNN	This paper introduces a text classification method that improves accuracy, precision, recall, and F1-score in natural language processing by integrating CNN, RNN, and BiLSTM models.	Traditional natural language processing methods such as Word2Vec, CNN, RNN, and BiLSTM often struggle to capture nuanced contextual information.
3	2023	Ensemble Stacking Model for Sentiment Analysis of Emirati and Arabic Dialects	RNN, LSTM, CNN	The study develops a deep learning model achieving high accuracy in sentiment analysis of Emirati Arabic dialect using four distinct datasets.	Achieving an impressive accuracy rate of 98.53%, the study introduces a deep learning model tailored for sentiment analysis in both Arabic and Emirati dialects, spanning multiple datasets.
4	2022	A survey on sentiment analysis methods, applications, and challenges	NLP, LSTM, RNN, Bi-RNN,ANN	This article encompasses the evolution of web-based applications, sentiment analysis, its various applications, evaluations, encountered challenges, as well as the benefits and drawbacks associated with sentiment analysis.	This article compares machine learning, neural networks, and SVM algorithms for sentiment analysis, emphasizing features like feature selection, data collection, and application domains.
5	2023	A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges	RNN, Seq2Seq,LSTM, DL	This survey explores aspect-based sentiment analysis (ABSA), including recently developed compound tasks, pre-trained language models, and useful ABSA systems.	Aspect-based sentiment analysis (ABSA) is explored in this survey, covering pre-trained language models, useful ABSA systems, and recent developments in compound tasks.
6	2022	Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion	NLP, MSA, AI, CNN	In the context of NLP and AI, this paper examines Multimodal Sentiment Analysis (MSA), contrasting its development, outlining its benefits and drawbacks, and suggesting future directions.	MSA research uses fusion techniques, bimodal modalities, context-aware humor detection, and application-specific architecture modifications to improve emotion recognition and sentiment prediction, mainly at the word level.



		methods, applications, challenges and future directions			
7	2022	A systematic review of social media-based sentiment analysis: Emerging trends and challenges	CNN, LSTM, Bi-LSTM, KNN	The PRISMA framework is used to analyze sentiment in social media, highlighting its limitations like the lack of English-language publications and difficulty in understanding non-English text.	This paper compares studies, examines challenges and potential pitfalls in the investigation of sentiment analysis in social media, and identifies problems with datasets, text languages, and analysis techniques.
8	2023	Survey on sentiment analysis: evolution of research methods and topics	LSTM, CNN, DNN, DBN, RNN	The study employs the PRISMA framework to analyze sentiment in social media, focusing on user-generated data and associated challenges. It identifies limitations and provides recommendations for future research endeavors.	This study uses LSTM to improve the processing speed of the hidden layer posterior of the RNN module for sentiment analysis.
9	2023	Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research	NLP, AI, LSTM, CNN	Through the utilization of keyword co-occurrence analysis and a community detection algorithm, the study conducts sentiment analysis, identifies trends, provides guidance, and directs technical direction.	By examining user comments, social media platforms, sentiment analysis methods, and non-English language analysis using informetric tools and keyword co-occurrence analysis, this paper explores the development of sentiment analysis research methods.
10	2023	Multimodal Sentiment Analysis Based on Fusion Methods: A Survey	SVM, LSTM, CNN, MKL, THMM	This paper investigates the evolution of sentiment analysis research methods by analyzing user comments, social media platforms, sentiment analysis techniques, and non-English language analysis through informetric tools and keyword co-occurrence analysis.	The paper highlights current developments in multimodal sentiment analysis and makes recommendations for future directions in the field, with a particular emphasis on conversational sentiment. It emphasizes how subjectivity analysis and entity recognition could be improved by these tools.
11	2023	More than a Feeling: Accuracy and Application of Sentiment Analysis	ML, SVM, TL	Sentiment polarity can be effectively inferred through the use of sentiment analysis, a technology that combines audio and visual data. Its uses include evaluating product reviews, analyzing public opinion, and assessing risk.	This study looks at how sentiment analysis changed between 2012 and 2022 thanks to advances in artificial intelligence. It emphasizes how it might be used to improve customer experience and conduct competitive strategy research.
12	2023	Restaurant survival prediction using	GCN, CNN, LSTM, DL	Aspect-based sentiment analysis is used in the study to forecast restaurant survival in popular US tourist locations.	The study shows that the prediction performance of restaurant survival is improved by sentiment analysis of online reviews. It highlights how



		customer-generated content: An aspect-based sentiment analysis of online reviews		The prediction performance is improved through the identification of the most influential features found in online reviews.	important location and food quality are as key components in using the CSF model to validate user-generated content.
13	2022	A Large-Scale Group Decision Making System Based on Sentiment Analysis Cluster	NLP, classification, clusters, SVM	This paper presents a novel approach to large-scale group decision making that combines bag-of-words and sentiment analysis to handle data produced by multiple experts. Conducting behavior studies in large settings is made easier with this method.	The paper presents a novel LSGDM system to improve Group Decision Making (GDM). This system uses sentiment analysis to assess the degree of positivity and aggression in expert comments, allowing consensus weights to be adjusted and groups to be created.
14	2023	A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research	NLP, DL, SVM, Naive Bayes, Logistic Regression, Random Forest	Text is categorized using sentiment analysis into three categories: neutral, negative, and positive. Research highlights the significance of this method for political analysis, banking, and organizations while also pointing out obstacles, opportunities, and developments.	Given its susceptibility to poorly constructed texts, sentiment analysis becomes an important area of research in natural language processing. It uses methods from deep learning to guide strategic decision-making.

II. PROPOSED SYSTEM

In the rapidly growing technologies in IT sector, sentimental analysis has very large scope in future in perspective of customer as well as Businessman. There are many ways of performing sentimental analysis such as video analysis, text analysis, text summarization, and many more. There are many sentiments analysis project going on but the problem they are facing some problem like it will give the question or some type of questions which will define user’s behavior. In this paper is cost-efficient, easy to use the website for any new on, requires less time to give the analytical report.

In this paper, developing a sentiment analysis website that is very easy to use and eco-friendly. In this paper we are creating a website in which we are feeding raw data as input and report is based on the input. There are different deep learning libraries that ca be used for making a website integrated sentimental analysis.[1] The difference that are adding to this is that there are applications but creating a website will be accessible to all the users. Plus, there is website summarization meaning it can analysis the website and give the sentimental analysis report accordingly.

Let’s get to the system and understand the concept that had used:

- o Hardware Requirements
- o Software Requirements

1. Hardware Requirements’:

- o Core i5/i7 processor
- o At least 8 GB RAM
- o At least 60 GB of Usable Hard Disk Space

2. Software Requirements

In website integrated sentimental analysis development, using deep learning libraries like NLTK, LSTM and .Json, Html, JavaScript for website.

At the first created a website that shows the sentimental analysis present in the Text feed in the model. It will show if the text is positive, negative or neutral. This will help the user and to know the perspective of a product if it is good or not. This will reduce the time that will be required to read all the reviews of the product. Got the idea of creating a website and Extension instead of application.[2]

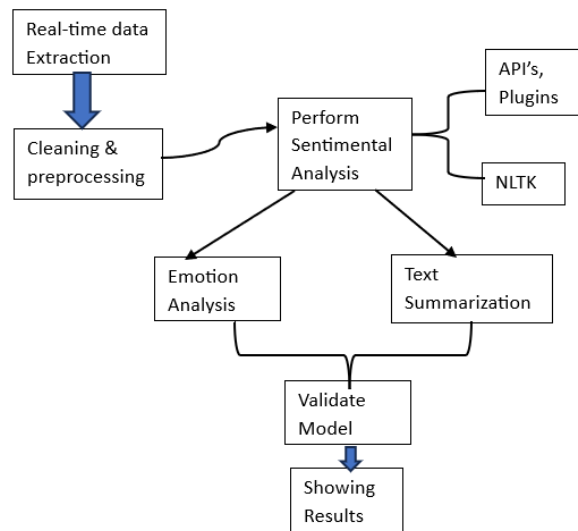


Fig -1 Proposed System

In figure No.1, taking the real time data for interpretation and evaluating the website, text and image and preparing a report based on the input fed in the system. Let's look more in details:

1. Real- Time Data Extraction:

In this step, taking the real time data from the user and extracts the valuable data from the website or text for next process.

2. Cleaning & Preprocessing:

In this, the input fed by the user is cleaned and pre-processed because the system focuses on the sentiment not all the data fed in the system. It will choose which word shows sentiment in the website or text.

3. Perform Sentimental Analysis

To perform this step, there are 4 steps that should be performed:

API's Plugins:

In this data is connected to API's and Plugins that will help the website to take inputs for model developing.

NLTK:

Using deep learning libraries like NLP, NLTK, etc. to take input data. In this NLTK libraries and python are used for text processing. It will help the process as the libraries have natural language feature that will read the text and analysis it. [6][9]

Text Summarization:

Using NLTK libraries, Text summarization is given by the model. That will provide a summary that will indicate the sentiments of the text.

Emotion Analysis:

In this step, Text is classified into 3 class that is Positive, Negative and Neutral. From this, user can check the sentiments of the text and provide the report of sentimental analysis.

Validate Model:

After getting the result of the emotions analysis, model will check the dataset that how much the model shows accuracy in which class like positive, negative, and neutral.

Showing Results:

In this step, gathering all the data and reports or output form the above steps and showing the reports to the users about the sentiments that are in the text that was provided by the user. Showing the accuracy of the sentiments that will help the user to know about the Text that cab be skipped sometimes. This model will help the user to know about the sentiments present in the Text.

III. ALGORITHM

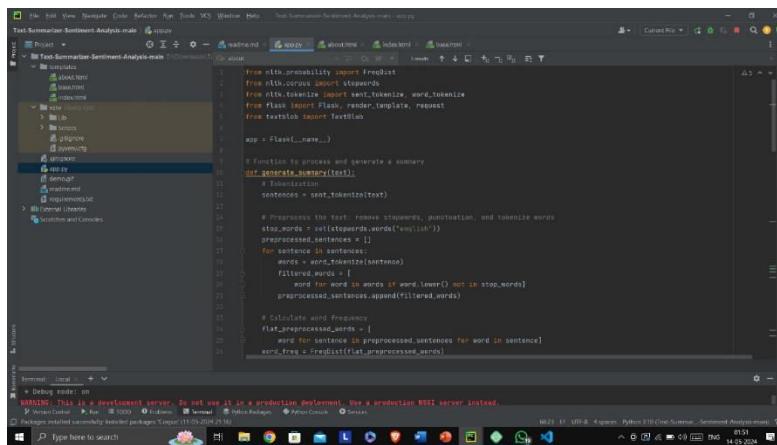
In this paper, following steps are followed for creating the extension and website for sentimental analysis.

o Text Preprocessing:

This step involves 2 process:

- Tokenization:

The segmentation of the input text into sentences is facilitated by the sent_tokenize() function from the NLTK library, while the word_tokenize() function is employed to tokenize each sentence into individual words.



```

from nltk.probability import FreqDist
from nltk.corpus import stopwords
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk import FreqDist, PorterStemmer, TextBlob
from textblob import TextBlob

stop = FreqDist(['_'])

# Function to process and generate a summary
def generate_summary(text):
    # Tokenization
    sentences = sent_tokenize(text)

    # Preprocess the text: remove stopwords, punctuation, and tokenize words
    stop_words = set(stopwords.words('english'))
    preprocessed_sentences = []
    for sentence in sentences:
        words = word_tokenize(sentence)
        filtered_words = [
            word for word in words if word.lower() not in stop_words
        ]
        preprocessed_sentences.append(filtered_words)

    # Calculate word frequency
    flat_preprocessed_words = []
    for sentence in preprocessed_sentences:
        for word in sentence:
            flat_preprocessed_words.append(word)

    word_freq = FreqDist(flat_preprocessed_words)
    
```

Fig -2 Code for Text Preprocessing

- Stop-word Removal

The NLTK library's stopwords.words("english") function furnishes a list of common English stopwords, which are words that typically convey little semantic value. By eliminating these stopwords from the tokenized words, noise in the text data is mitigated.

o Generate Summary

The summarization algorithm operates in several steps:

- Word Frequency Calculation:

Utilize NLTK's FreqDist() function to compute the frequency of each word in the preprocessed text.

- Sentence Scoring:

Score each sentence by summing the frequencies of the words it contains.

- Iterate over each tokenized sentence:

For each word in the sentence, retrieve its frequency from the FreqDist object. Sum the frequencies of all words in the sentence to obtain the sentence score.

- Top Sentences Selection:

Select the top-ranked sentences, typically the top 3 sentences in this case, to form the summary.

o Text Sentiment Analysis:

Sentiment analysis is conducted on the original text using the TextBlob library. The TextBlob() class is employed to create a TextBlob object, which offers methods for sentiment analysis, including sentiment.polarity. This method returns a polarity score indicating the positivity or negativity of the text.

The sentiment is categorized based on the polarity score. If the score exceeds 0.1, the text is classified as positive. Conversely, if it falls below -0.1, it is categorized as negative. Any score within the range of -0.1 to 0.1 is considered neutral.

o Summary Sentiment Analysis:

Just like the sentiment analysis conducted on the original text, sentiment analysis is also performed on the generated summary using TextBlob.

The generated summary is converted into a TextBlob object, and its sentiment polarity score is calculated using the sentiment.polarity method provided by TextBlob.

Similar to the sentiment categorization process for the original text, the sentiment polarity score of the summary is categorized as positive, negative, or neutral using the same threshold values. If the score exceeds 0.1, the summary is labeled as positive, while a score below -0.1 indicates a negative sentiment. Any score within the range of -0.1 to 0.1 is considered neutral.

These all algorithms work together in flask application to process the input data, generate summary and analyze and give sentimental Analysis Report that will provide insight to the user.

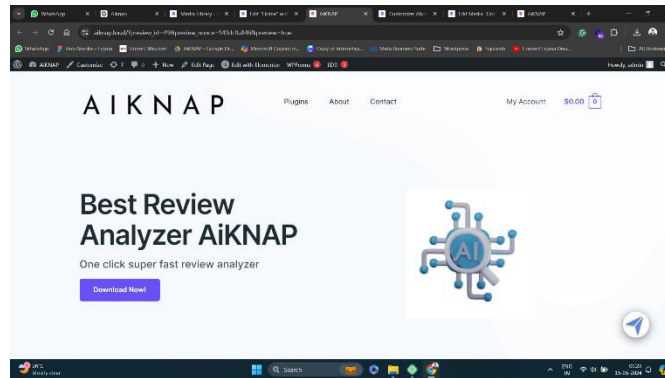


Fig -3 Website (Dashboard)

In figure no.2, this includes the website overview or dashboard that will help to download the extension of the model. So that it can be accessible to all the websites and show the sentiment analysis. It also incorporates chatbots so that the new users don't face difficulties while using it.

IV. RESULT & CONCLUSION

In Previous years, there are many sentimental analyses for the text but that is in application and can cost or charges to buy premium version of the application.[4] Instead of this, creating an extension that will be opensource for everyone. It will also work on Deep learning libraries that will enhance the ability to analyses the text and give the accurate sentimental analysis report. [6]

In this work, using NLP, NLTK of deep learning libraries that will provide sentimental analysis of the text. The main objective of this work includes amalgamating computer vision and NLP methodologies to create an online website with extension. It will be capable of categorizing data into positive, negative, and neutral sentiments. It provides the insights that will help the user for understanding the sentimental analysis in the text. This work enhances the business decision making, improve customer satisfaction, efficient data analysis and holistic website and extension.[13]

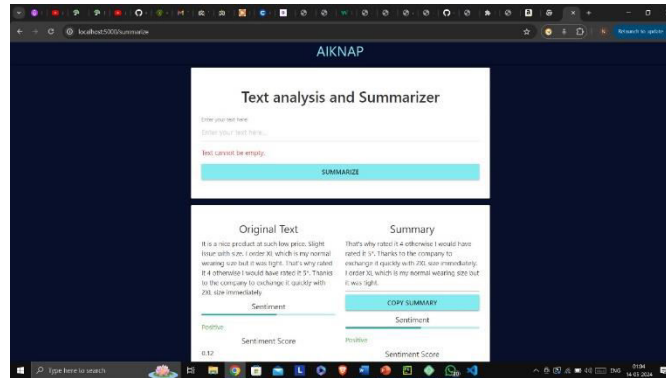


Fig -4 Text Analysis & summarization

In figure no.4, input data is fed to the model then it will give sentiments present in the text in form of positive, negative, and neutral.

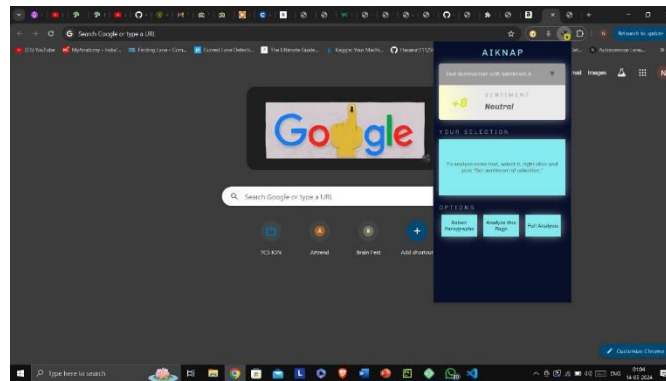


Fig -5 Extension of the Website

In figure no.5, an extension of the website is downloaded from the dashboard. And it is showing the sentimental analysis of the website of google. Since, the google home doesn't contain any kind of emotion. So it is giving neutral as sentimental analysis.

The suggested sentiment analysis system offers a complete solution for companies and individuals looking for textual data insights by fusing deep learning techniques into user-friendly applications.

In an essence, proposed aiknap model performed well in website analysis and also in extension, able to analyse the sentiments in the text and website.

REFERENCES

1. Ankita Gandhi, Kinjal Adhvaryu, Soujanya Poria, Erik Cambria, Amir Hussain, "Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions" in 2022
2. Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam," A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 35, NO. 11, NOVEMBER 2023
3. Arwa Alqarni and Atta Rahman," Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach", in 2023
4. Kian Long Tan, Chin Poo Lee and Kian Ming Lim, "A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research", in 2023
5. Arwa A. Al Shamsi, Sherief Abdallah, "Ensemble Stacking Model for Sentiment Analysis of Emirati and Arabic Dialects", in 2023

6. Abayomi Bello, Sin-Chun Ng and Man-Fai Leung, “A BERT Framework to Sentiment Analysis of Tweets”, in 2023
7. Hengyun Lia, Bruce X.B. Yub, Gang Lic, Huicai Gaoa, “Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews”, in 2023
8. Linan Zhua, Zhechao Zhua, Chenwei Zhangb, Yifei Xua, Xiangjie Konga, “Multimodal Sentiment Analysis Based on Fusion Methods: A Survey”, in 2023
9. Victor Chang, Lian Liu, Qianwen Xu, Taiyu Li, Ching-Hsien Hsu, “An Improved Model for Sentiment Analysis on Luxury Hotel Review”, in 2023
10. Jos´e Ram´on Trilloa, Enrique Herrera-Viedmaa, Juan Antonio Morente-Molineraa, Francisco Javier Cabrerizo “A Large-Scale Group Decision Making System Based on Sentiment Analysis Cluster”, in 2022.
11. Jochen Hartmann, Mark Heitmann, Christian Siebert, Christina Schamp, “More than a Feeling: Accuracy and Application of Sentiment Analysis” in 2022.
12. Hamed Taherdoost, Mitra Madanchian, “Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research” in 2023
13. Jingfeng Cui, Zhaoxia Wang, Seng Beng Ho, Erik Cambria, “Survey on sentiment analysis: evolution of research methods and topics” in 2023
14. Marco A. Palomino, Farida Aider, “Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis” in 2022
15. Qianwen Ariel Xu, Victor Chang Chrisina Jayne, “A systematic review of social media-based sentiment analysis: Emerging trends and challenges,” in 2022
16. Mayur Wankhade, Annavarapu Chandra Sekhara Rao, Chaitanya Kulkarni, “A survey on sentiment analysis methods, applications, and challenges,” in 2022



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details