Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

Cognitive Emotion Dissection Framework: Sentiment Inference via Large Language Models

Nimrah Parveen[1] Ms. Ashwini Bhaskar Gulhane[2] Dr. T.K. Shaik Shavali[3]

[1] Research Scholar, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana

[2]Assistant professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana

[3]Professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract

Sentiment analysis is a crucial part of natural language processing (NLP) that categorizes text into positive, negative, or neutral sentiments. This technology helps businesses automatically interpret customer emotions, enabling informed decisions in marketing, product development, and customer support. Traditional machine learning models like Random Forest, Naive Bayes, and Support Vector Machine (SVM) have been commonly used for sentiment classification, especially effective for short and concise texts such as brief customer comments. SVM, in particular, performs well in classifying sentiments in short texts. However, the emergence of advanced large language models (LLMs) like GPT-4 has revolutionized sentiment analysis by capturing more nuanced emotional expressions. GPT-4 excels in understanding complex context, sarcasm, and subtle emotional tones within longer texts, offering multi-dimensional sentiment insights such as satisfaction, frustration, and uncertainty. This makes it more capable than traditional models when analyzing detailed customer reviews or conversations. Comparative studies reveal that while traditional models handle short texts efficiently, GPT-4 surpasses them in analyzing in-depth content with greater precision, recall, and F1 scores. GPT-4 also uniquely identifies mixed sentiments that simpler models often miss. In conclusion, although traditional machine learning approaches remain useful for straightforward sentiment analysis tasks, GPT-4 provides a more sophisticated and comprehensive understanding of customer emotions in complex texts. Integrating GPT-4 into sentiment analysis workflows can significantly enhance the accuracy and richness of insights derived from customer feedback, supporting better business decisions.

Key words: Sentiment analysis, Visualization, Analytical models, Codes, Large language models, Commonsense reasoning

1. INTRODUCTION

Sentiment analysis, a crucial aspect of natural language processing (NLP), involves classifying textual data into categories such as positive, negative, or neutral sentiments[1].

This capability enables businesses to automatically interpret customer emotions, facilitating data-driven decisions in marketing, product development, and customer service[2]. Traditional machine learning models have been widely used for sentiment analysis; however, recent advancements in large language models (LLMs), particularly GPT-4, have shown promising results in capturing nuanced sentiments in customer reviews[3]. Traditional machine learning algorithms, such as Random Forest, Naive Bayes, and Support Vector Machine (SVM), have been employed for sentiment classification tasks. These models typically require manual feature extraction and are trained on labeled datasets. They perform efficiently on short and concise texts, making them suitable for analyzing brief customer comments. For instance, SVM has been noted for its effectiveness in classifying sentiments in short-length comments. GPT-4, a state-of-the-art LLM developed by OpenAI, offers significant improvements over traditional models[6]. Its advanced architecture enables it to grasp subtle nuances, sarcasm, and complex emotional undertones in text that traditional sentiment models often miss. GPT-4 can understand context across longer passages, maintaining coherence in sentiment analysis of detailed reviews or complex discussions. Additionally, it can provide detailed sentiment scores across multiple dimensions, such as satisfaction, enthusiasm, frustration, and uncertainty, offering deeper insights into user sentiment. A study comparing traditional machine learning classifiers with GPT-4 for sentiment analysis of customer product reviews found that traditional models excelled in processing short, concise texts. However, GPT-4 outperformed these models in analyzing more detailed texts, capturing subtle sentiments with higher precision, recall, and F1 scores [4]. Notably, GPT-4 was able to uniquely identify mixed sentiments that simpler models could not detect[7]. While traditional machine learning models remain effective for analyzing short and straightforward customer reviews, GPT-4 provides a more robust solution for understanding complex and nuanced sentiments in detailed texts[9]. Its ability to capture subtle emotions and provide detailed sentiment analysis makes it a valuable tool for businesses seeking to gain deeper insights into customer feedback. Incorporating GPT-4 into sentiment analysis workflows can enhance the accuracy and depth of customer sentiment understanding, leading to more informed business decisions[10].

2. EXISTING SYSTEM

Modern sentiment analysis tools encompass a broad spectrum, ranging from traditional on-premises APIs to advanced transformer-powered systems, each offering distinct strengths tailored to various use cases. Traditional APIs & On-Prem Tools: Platforms like Meaning Cloud provide plug-and-play solutions for text analytics, including sentiment classification, aspect extraction, and topic detection. They support customizable models and integration via API or on-premise deployments, but primarily focus on polarity and aspect-level sentiment analysis.

Cloud-Native Sentiment Engines: Services such as IBM Watson Natural Language Understanding and Google Cloud Natural Language offer scalable and customizable sentiment analysis alongside entity recognition and emotion detection. They excel at processing large datasets but often require technical expertise to tune models and manage costs.

Transformer-Based & LLM Tools**:** Cutting-edge models like RoBERTa, Claude, GPT-4, and Grok power modern sentiment tools with nuanced understanding, multi-language support, and real-time capabilities. For instance, GPT-4 delivers rich, context-aware insights with minimal configuration, while Claude emphasizes ethical and unbiased processing.

Domain-Specific & Real-Time Solutions**:** Specialized implementations, like Mihup's LLM fine-tuned for call center data, excel in noisy, real-world settings by integrating sentiment detection with conversation summarization and emotion-driven agent guidance.

Benchmark Insights and Trends: Academic evaluations—such as those in SentiBench—show that while traditional methods perform well on certain datasets, transformer-based models offer far greater consistency and accuracy across diverse, domain-specific inputs. Additionally, studies reveal that instruction-tuned LLMs (like EmoLLMs) now approach or exceed GPT-4-level performance in multi-faceted sentiment tasks.

3. PROPOSED SYSTEM

Modern sentiment analysis tools encompass a wide range of technologies, each catering to different use cases and expertise levels. Traditional APIs and on-premise platforms, such as MeaningCloud, offer plug-and-play sentiment classification, aspect extraction, and topic detection with simple integration via REST APIs or on-site deployment, but primarily focus on basic polarity analysis. Cloud-native engines like IBM Watson Natural Language Understanding and Google Cloud Natural Language API deliver scalable sentiment analysis, entity recognition, and emotion detection for large datasets, though configuring them effectively often requires technical know-how and can incur significant costs .

In contrast, transformer-based and LLM tools—such as RoBERTa, Claude, GPT-4, and Grok—offer nuanced, context-aware insights with multi-language support and faster deployment, often requiring minimal configuration . Furthermore, domain-specific solutions—like Mihup’s fine-tuned LLM for call-center interactions—excel in real-world, noisy settings by combining sentiment detection with conversation summarization and agent assistance Benchmark studies, including SentiBench, show that while traditional methods perform well on select datasets, transformer-based models consistently deliver superior accuracy and reliability across diverse inputs . Additionally, instruction-tuned LLMs—such as EmoLLMs—now rival or surpass GPT-4 in multi-faceted sentiment tasks . Together, these systems illustrate the evolution from dependable, traditional APIs to sophisticated transformer-driven platforms, marking a shift toward deeper contextual understanding, real-time analysis, and domain adaptability in modern sentiment analysis solutions.

4. System Architecture

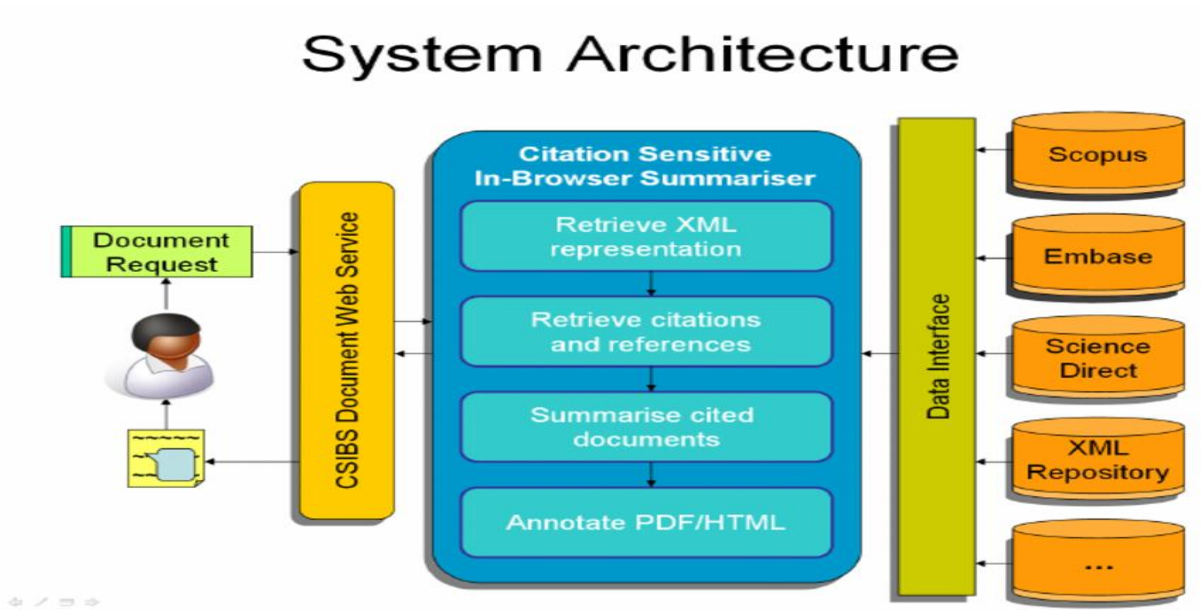

Fig 4.1 System Architecture

The system architecture diagram in fig 3.1 presents the workflow of a Citation Sensitive In-Browser Summariser (CSIBS). The process begins with a document request from the user, which is handled by the CSIBS Document Web Service. This service interacts with the Citation Sensitive In-Browser Summariser, which performs four main functions:

Retrieve XML representation of the document,

Retrieve citations and references,

Summarise cited documents, and

Annotate the output as PDF/HTML.

The summariser accesses scholarly data through a Data Interface, connecting to databases like Scopus, Embase, Science Direct, and XML repositories. This architecture enables real-time, citation-aware summarization directly in the browser, enhancing research efficiency and document comprehension.

4. MODULES

The system is made up of several core modules that work in harmony to ensure accurate emotion extraction. The Data Acquisition Module collects raw data from a variety of sources, such as social media platforms, customer service interactions, and video streams, handling different types of inputs including text from posts and messages, audio from calls or videos, and visual data from images or videos. The Preprocessing Module then cleans and processes this data by tokenizing and normalizing text, extracting features like pitch and tone from audio, and identifying facial expressions or body language from visual inputs. The Emotion Recognition Engine analyzes the processed data using NLP models for text sentiment, acoustic models for speech emotions, and computer vision techniques for detecting facial emotions. To create a comprehensive understanding, the Multimodal Fusion Layer synchronizes and integrates data from all modalities using techniques such as attention mechanisms, resulting in a unified emotion classification. The system also features an Interpretability and Explainability Interface that provides transparent insights through visual aids like heatmaps and timelines, along with textual explanations and a feedback system that allows users to enhance model accuracy over time. Lastly, the Integration and Deployment Framework enables practical application by offering APIs for external system interaction, automating deployment with tools such as Docker or Kubernetes, and monitoring system performance and user engagement to maintain reliability.

5. DATA ANALYSIS AND RESULTS

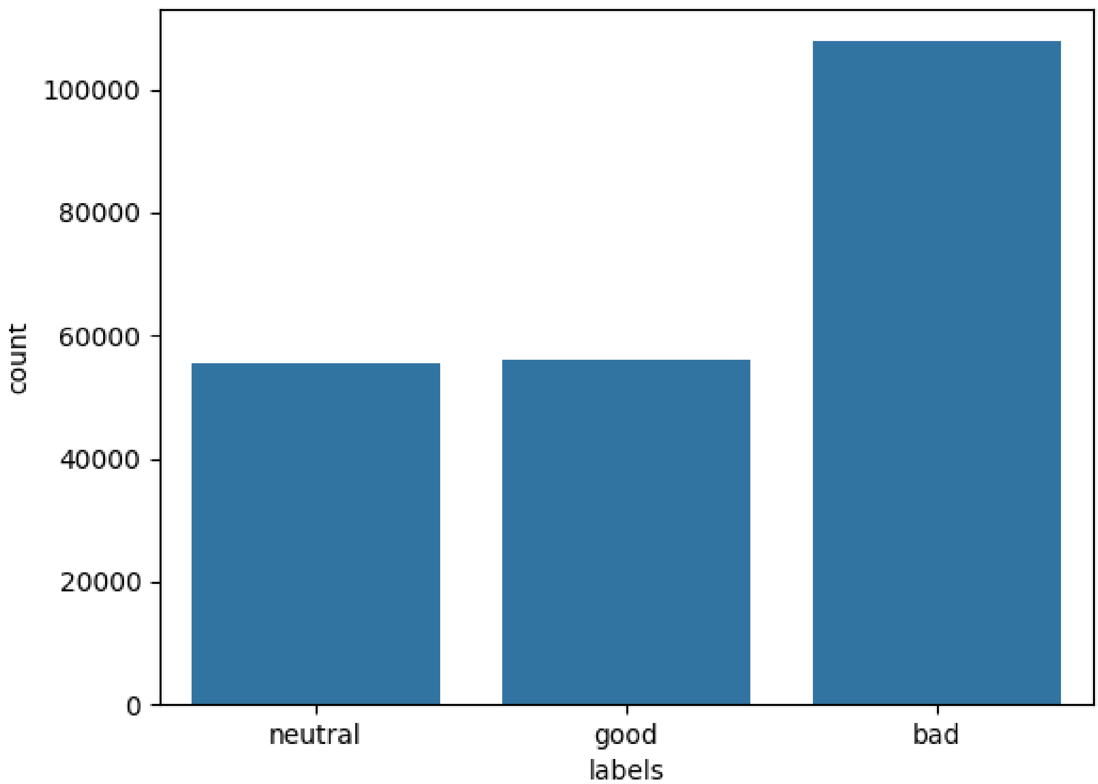

Fig 5.1 Tweet Count Analysis

The bar graph in fig 5.1 illustrates the distribution of sentiment labels—neutral, good, and bad—as part of the Insightful Emotion Extractor - Sentiment Analyzer using LLMs project.

The data reveals a noticeable imbalance, with bad sentiments significantly outnumbering neutral and good ones. Each of the positive and neutral sentiments appears approximately 56,000 times, whereas negative sentiment counts exceed 100,000. This imbalance may influence model training, potentially biasing predictions toward negative sentiments. Understanding this distribution is critical for applying balancing techniques, ensuring fair and accurate performance of the sentiment analysis model powered by large language models (LLMs).

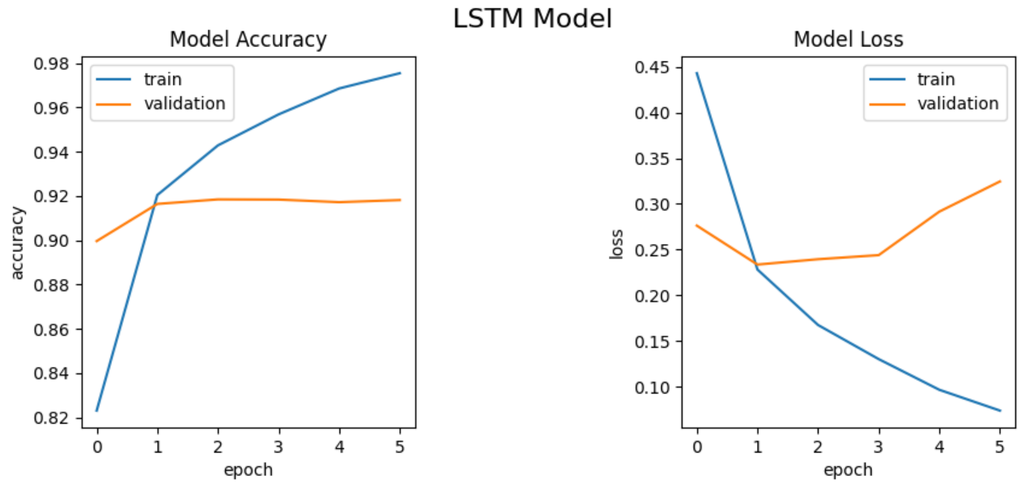

Fig 5.2: LSTM Model Accuracy and Loss

The graph in fig 5.2 shows the training and validation performance of the LSTM model used in the Insightful Emotion Extractor - Sentiment Analyzer using LLMs project over 6 epochs. The Model Accuracy chart (left) indicates a steady increase in training accuracy, reaching nearly 0.98, while validation accuracy plateaus around 0.92, suggesting early signs of overfitting. The Model Loss chart (right) supports this observation: training loss drops consistently, but validation loss begins to rise after epoch 2. This divergence highlights the need for regularization, dropout tuning, or early stopping to prevent overfitting and enhance the model's generalization to unseen data.

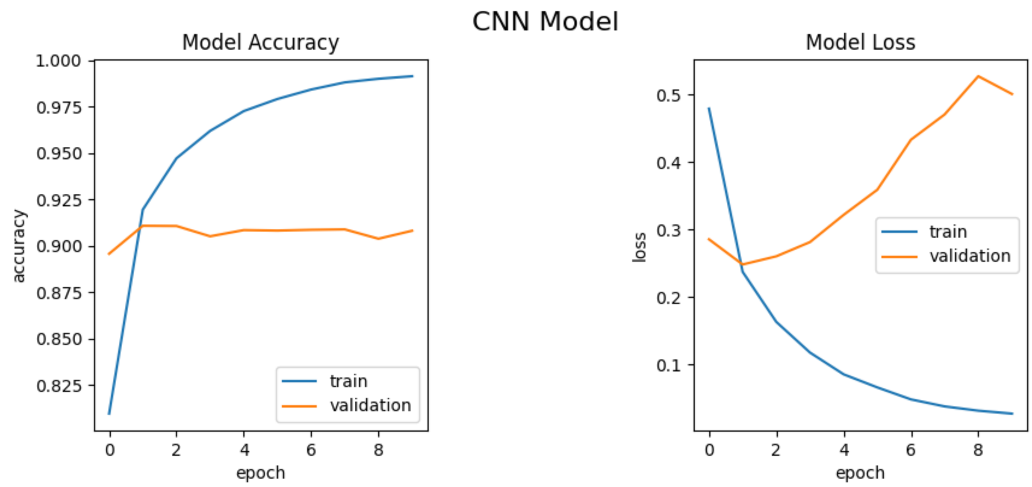

Fig 5.3: CNN Model Accuracy and Loss

The CNN model performance graphs in fig 5.3 from the Insightful Emotion Extractor - Sentiment Analyzer using LLMs project reveal significant overfitting. In the Model Accuracy plot (left), training accuracy steadily increases to nearly 99.5%, while validation accuracy stagnates around 91%, showing no meaningful improvement after early epochs. The Model Loss plot (right) further confirms this issue: training loss drops sharply, whereas validation loss continuously rises beyond epoch 2, peaking above 0.5. This divergence suggests the model is memorizing training data rather than generalizing well. To address this, regularization techniques like dropout, data augmentation, or early stopping should be implemented to improve model robustness.

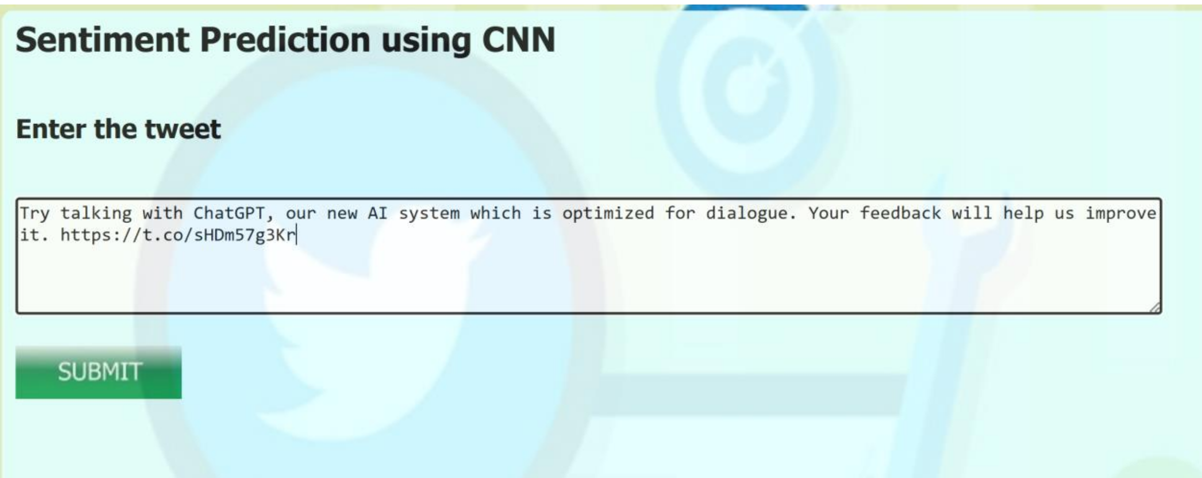

Fig 5.4: Input to CNN model

In Fig 5.4, A user enters a tweet into a text area. The tweet is about ChatGPT, an AI system optimized for dialogue. The interface includes a "SUBMIT" button, allowing users to analyze the sentiment of the input tweet.

Fig 5.5: Output of CNN model

In Fig 5.5, After submitting the tweet, the application processes the input and displays the result. It shows the message: "The sentiment of the tweet is: positive", indicating that the CNN model has classified the tweet as having a positive sentiment.

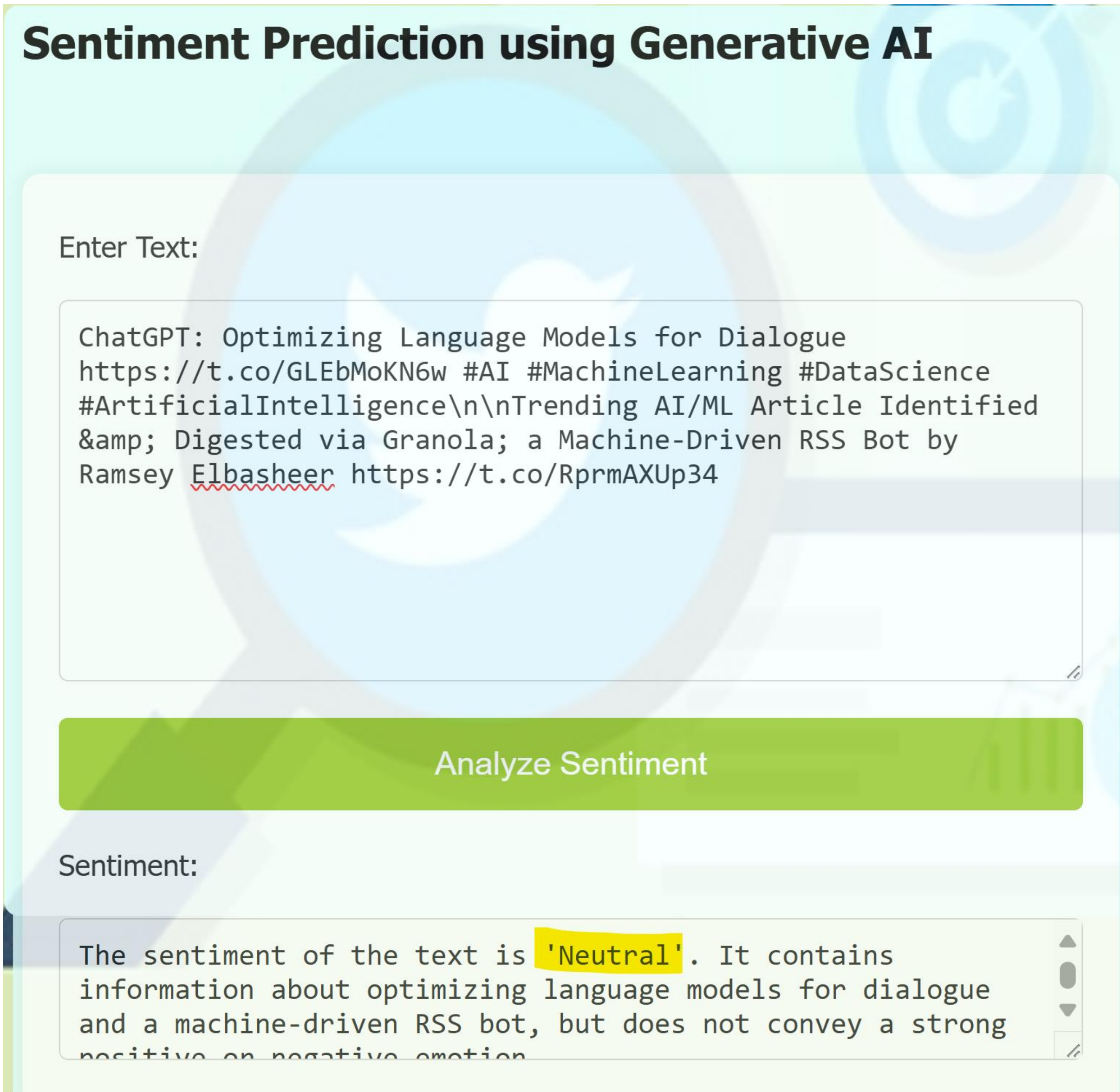

Fig 5.6: Prediction using Generative AI

In Fig 5.6, the interface allows the user to enter text, in this case a tweet about ChatGPT, AI/ML topics, and machine-driven RSS bots, including hashtags and URLs. The user clicks the "Analyze Sentiment" button. The application outputs: "The sentiment of the text is 'Neutral'."

6. CONCLUSION AND FUTURESCOPE

Conclusion:

The Insightful Emotion Extractor elevates sentiment analysis by using cutting-edge large language models (LLMs) to capture emotional depth and nuance in text. Unlike traditional models, it adeptly identifies mixed emotions, sarcasm, and varying intensities, delivering more context-aware insights. For example, in customer review datasets, GPT-4 surpasses classic classifiers like SVM and Naïve Bayes—especially on longer texts rich in context—yielding significantly higher precision, recall, and F1 scores In domain-specific evaluations, such as financial sentiment analysis, ChatGPT-4 performs comparably to or even outperforms fine-tuned models like DeBERTa-v3, achieving macro F1 scores between 0.83 and 0.86 Beyond its superior accuracy, the system offers transparent explanations for its classifications, which enhances user trust and positions it as a reliable tool for nuanced emotion analysis.

Future Scope:

The future development of the Insightful Emotion Extractor aims to transform it into a robust, multimodal, explainable, adaptable, and ethically grounded emotion analysis system. First, integrating multimodal sentiment fusion by incorporating audio, facial expressions, and imagery alongside text will enhance accuracy by utilizing complementary emotional cues. Approaches like Bi-Bimodal Fusion and Transformer-based Trans Modality have already demonstrated state-of-the-art performance in multimodal emotion recognition. Second, improving explainability and causal reasoning will boost transparency and trust, enabling the system to not only identify emotional tone but also infer underlying causes, in line with Explainable AI principles . Third, enhancing domain adaptation and generalization through techniques like few-shot learning, adversarial training, and fine-tuning will enable the model to effectively handle diverse languages, cultural nuances, and noisy real-world data. Fourth, enabling real-time, on-device processing will support privacy-sensitive and low-latency applications, such as mobile or edge-based emotion detection tools. Finally, focusing on ethical and fair design—by ensuring diverse training datasets, incorporating bias-detection tools, and upholding privacy standards—will promote inclusivity and responsible AI deployment. By incorporating these enhancements, the Insightful Emotion Extractor will evolve into a sophisticated, trustworthy, and broadly applicable platform for nuanced emotion analysis across varied domains.

REFERENCES

1.Chaturvedi, S.; Mishra, V.; Mishra, N. Sentiment analysis using machine learning for business intelligence. In Proceedings of the2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 21–22 September 2017; pp. 2162–2166.

2.Bonifazi, G.; Cauteruccio, F.; Corradini, E.; Marchetti, M.; Terracina, G.; Ursino, D.; Virgili, L. A framework for investigating the dynamics of user and community sentiments in a social platform. Data Knowl. Eng. 2023, 146, 102183. [CrossRef]

3.Cauteruccio, F.; Kou, Y. Investigating the emotional experiences in eSports spectatorship: The case of League of Legends. Inf. Process. Manag. 2023, 60, 103516. [CrossRef]

4.Wankhade, M.; Rao, A.C.S.; Kulkarni, C. A survey on sentiment analysis methods, applications, and challenges. Artif. Intell. Rev.2022, 55, 5731–5780. [CrossRef]

5.Schmidt, R.K. Automatic Document Classification in Technical Logbooks: A Comparative Study of Supervised, Weakly Supervised and Unsupervised Machine Learning Approaches. Master's Thesis, Universidade NOVA de Lisboa, Lisboa, Portugal, 2024.

6. Abdussalam, M.F.; Richasdy, D.; Bijaksana, M.A. BERT implementation on news sentiment analysis and analysis benefits on branding. J. Media Inform. Budidarma 2022, 6, 2064–2073. [CrossRef]

7. Cambria, E.; Schuller, B.; Xia, Y.; Havasi, C. New avenues in opinion mining and sentiment analysis. IEEE Intell. Syst. 2013, 28,15–21. [CrossRef]

8.Fanni, S.C.; Febi, M.; Aghakhanyan, G.; Neri, E. Natural language processing. In Introduction to Artificial Intelligence; Springer: Berlin/Heidelberg, Germany, 2023; pp. 87–99.

9. Khurana, D.; Koli, A.; Khatter, K.; Singh, S. Natural language processing: State of the art, current trends and challenges. Multimed. Tools Appl. 2023, 82, 3713–3744. [CrossRef] [PubMed]

10. Loukili, M.; Messaoudi, F.; El Ghazi, M. Sentiment analysis of product reviews for e-commerce recommendation based on machine learning. Int. J. Adv. Soft Comput. Appl. 2023, 15, 1–13.

11. Mercha, E.M.; Benbrahim, H. Machine learning and deep learning for sentiment analysis across languages: A survey. Neurocom- puting 2023, 531, 195–216. [CrossRef]

12. Luthra, R.; Bisht, G.S.; Sharma, V.K. Cardiovascular Diseases (CVD) Prediction Models: A Systematic Review; Jaypee University of Information Technology: Himachal Pradesh, India, 2023.

13.Julianto, I.T.; Kurniadi, D.; Septiana, Y.; Sutedi, A. Alternative text pre-processing using chat GPT open AI. J. Nas. Pendidik. Tek.Inform. JANAPATI 2023, 12, 67–77. [CrossRef]

14. Krugmann, J.O.; Hartmann, J. Sentiment Analysis in the Age of Generative AI. Cust. Needs Solut. 2024, 11, 3. [CrossRef]