Explainable Sentiment Analysis for Social Media Post Using Attention Mechanisms

K. VIGNESH, P .VIHARI ,G.VIJAY KIRAN REDDY, P. VIJAY RAM REDDY,N.VIJAYALAKSHMI(students)

Prof. Dr.Sujith Das (Guide)

School of Engineering Department of AIML. Malla Reddy University, Hyderabad.

1.ABSTRACT:

The project was entitled " Explainable Sentiment Analysis for Social Media Posts Using Attention Mechanisms " the exponential growth of social media platforms has led to a surge in user generated content rich with opinions and sentiments, making sentiment analysis, a key aspect of natural language processing (NLP), essential for classifying sentiments into positive, negative, or neutral categories. This process has critical applications in marketing, customer feedback analysis, and political sentiment tracking. However, the unstructured nature of social media text, coupled with challenges like noise, misspellings, and irrelevant content, necessitates robust preprocessing and advanced analytical methods. This study leverages attention mechanisms to enhance sentiment analysis by improving prediction accuracy and interpretability. Attention mechanisms enable models to focus on the most relevant text elements, making decision-making more transparent. Using a dataset of millions of pre labeled tweets, the system integrates machine learning techniques like Logistic Regression and Multilayer Perceptron with Transformer-based architectures to achieve high classification accuracy. Dimensionality reduction through Principal Component Analysis (PCA) ensures computational efficiency without compromising performance

2.INTRODUCTION:

2.1 Problem Definition: The project enhances sentiment analysis on Twitter by improving accuracy, scalability, and interpretability. Twitter's unstructured text, including noise and misspellings, poses challenges for sentiment classification in marketing, customer feedback, and political analysis. This study leverages attention mechanisms to enhance prediction accuracy and transparency. Advanced machine learning models, including Logistic Regression, Multilayer Perceptron, and Transformer-based architectures, improve classification. Principal Component Analysis (PCA) ensures computational efficiency. Designed for large-scale Twitter datasets, the system provides precise, interpretable sentiment insights. By bridging the gap between model performance and user trust, it empowers organizations to make informed, data driven decisions**.**

2.2 Objective of Project: The project, "Explainable Sentiment Analysis for Social Media Posts Using Attention Mechanisms," focuses on enhancing sentiment analysis on Twitter by improving accuracy, interpretability, and scalability. It employs Logistic Regression, Multilayer Perceptron, and Transformer-based architectures to classify sentiments effectively. Attention mechanisms help models focus on key text elements, improving prediction transparency. Principal Component Analysis (PCA) ensures computational efficiency for large-scale Twitter data processing. The system handles unstructured text, addressing noise, slang, and misspellings. By delivering precise, explainable, and scalable sentiment analysis, the project provides actionable insights, empowering organizations to make data-driven decisions based on Twitter sentiment trends.

2.3 Scope & Limitations: Scope: The project aims to build a robust sentiment analysis system for Twitter, focusing on improving accuracy, interpretability, and scalability. Advanced machine learning techniques such as Logistic Regression, Multilayer Perceptron, and Transformer-based architectures will be employed to enhance sentiment classification. Attention mechanisms will be utilized to improve prediction transparency, allowing the model to focus on relevant text elements. The system is designed to handle 3 the unique challenges of Twitter data, including slang, emojis, and noise, providing more reliable and explainable sentiment insights.

3. LITERATURE REVIEW:

Sentiment analysis has been an area of extensive research in Natural Language Processing (NLP), with several studies focusing on improving accuracy, explainability, and efficiency. Over the years, researchers have developed various methodologies, ranging from lexicon-based approaches to deep learning models, to extract sentiment from text. This section reviews some of the most significant studies in sentiment analysis and their contributions.

3.1 Early Approaches: Lexicon-Based Sentiment Analysis

One of the earliest approaches to sentiment analysis was the lexicon-based method, where sentiment scores were assigned to words based on predefined dictionaries such as SentiWordNet and AFINN. Turney (2002) proposed an unsupervised learning approach that analyzed word co-occurrence patterns to classify reviews as positive or negative. This method worked well for structured datasets like product reviews but struggled with contextual nuances, sarcasm, and informal language found in social media.

In another study, Hu and Liu (2004) introduced opinion mining techniques, extracting sentiment from customer reviews using sentiment lexicons. While these methods were interpretable, they suffered from low accuracy due to the inability to handle contextual sentiment shifts

3.2 Machine Learning-Based Approaches

With the rise of machine learning, researchers developed supervised learning models to enhance sentiment classification. Pang et al. (2002) used Naïve Bayes, Support Vector Machines (SVM), and Maximum Entropy classifiers on movie reviews, achieving higher accuracy than lexicon-based methods. However, these models depended heavily on manual feature engineering, limiting their generalizability across different datasets.

In 2011, Socher et al. introduced Recursive Neural Networks (RNNs) for sentiment analysis, capturing contextual dependencies in text. Unlike traditional models, RNNs could analyze longer phrases, improving sentiment classification accuracy. Despite this advancement, RNNs faced vanishing gradient problems, limiting their performance on long-text sequences

3.3 Deep Learning and Transformer-Based Models

The introduction of deep learning and Transformer-based architectures revolutionized sentiment analysis. Kim (2014) proposed a Convolutional Neural Network (CNN) for sentence classification, demonstrating that CNNs could extract hierarchical text features without manual feature engineering. However, CNNs were less effective in capturing long-term dependencies in text.

A major breakthrough came with Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. Hochreiter and Schmidhuber (1997) developed LSTMs to address the vanishing gradient issue in RNNs, enabling better context retention. Wang et al. (2016) demonstrated that LSTMs outperformed traditional models in sentiment classification tasks. However, LSTMs still suffered from computational inefficiencies when processing large datasets.

The next leap in sentiment analysis came with Transformer-based models such as BERT (Bidirectional Encoder Representations from Transformers) by Devlin et al. (2018). BERT introduced contextualized word embeddings, capturing word meanings based on surrounding words. This significantly improved sentiment analysis performance, especially in handling sarcasm, negation, and complex sentence structures. RoBERTa, XLNet, and GPT models further enhanced sentiment classification by leveraging pre-trained deep learning architectures

Explainable Sentiment Analysis

Despite the success of deep learning models, their lack of explainability remained a challenge. Ribeiro et al. (2016) proposed LIME (Local Interpretable Model-Agnostic Explanations), allowing sentiment models to highlight important words influencing predictions. Recent studies, such as Xu et al. (2020), integrated attention mechanisms, improving model interpretability by visualizing sentiment contributions at the word level.

analysis has been a widely researched area in Natural Language Processing (NLP) for over two decades. Early studies focused on rule-based lexicon approaches, where predefined sentiment dictionaries like SentiWordNet and AFINN were used to assign polarity scores to words. While these methods worked well for structured text, they struggled with contextual meaning, sarcasm, and informal language present in social media data.

Study	Authors	Methodology Used	Findings
Sentiment Analysis Using Lexicon-Based Methods	Turney (2002)	PMI (Pointwise Mutual Information)	Struggles with context and sarcasm
Machine Learning for Sentiment Analysis	Pang et al. (2002)	Naive Bayes, SVM, Decision Trees	Improved accuracy but lacks contextual understanding
Deep Learning for Sentiment Analysis	Socher et al. (2013)	Recursive Neural Networks	Captures hierarchical word structures but computationally expensive
BERT for Sentiment Analysis	Devlin et al. (2018)	Transformer-based model (BERT)	State-of-the-art performance but lacks interpretability
Explainable AI for Sentiment Analysis	Ribeiro et al. (2016)	LIME (Local Interpretable Model-agnostic Explanatic	Improves interpretability of sentiment classification

4. Problem Statement:

Existing sentiment analysis models, including transformer-based approaches like BERT, provide predictions without clear justifications. Token-level explanations remain underexplored, making it difficult to understand how individual words influence.

Limited Utilization of Attention Mechanism: While attention mechanisms have been used for explainability, they are not fully leveraged for interpretability. The challenge is ensuring that attention reflects meaningful aspects of the text rather than arbitrary weight distributions.

Handling Informal Language, Slang, and Emojis: Social media language includes slang, code-mixing, and emojis, making sentiment analysis challenging. Many models struggle with such variations, leading to misclassification or loss of sentiment-related context.

This project aims to develop an Explainable Sentiment Analysis Model for Social Media Posts Using Attention Mechanisms to address these limitations. The model will provide token-level explanations to enhance transparency and interpretability. Advanced text preprocessing, domain-specific embeddings, and emoji sentiment mapping will improve sentiment detection in informal text. Bias-mitigation strategies will ensure fairness across different demographics. To enable real-time sentiment monitoring, lightweight architectures will be optimized for large-scale data streams. This model will support applications such as social media analytics, brand reputation management, and crisis detection, ensuring robust, explainable, and fair sentiment analysis.

5.Methodology:

The proposed sentiment analysis system aims to overcome the limitations of existing models by integrating explainability, bias mitigation, and real-time scalability using Transformer-based architectures with attention mechanisms. The system leverages Hierarchical Attention Networks (HAN), BERT-based models, and Explainable AI (XAI) techniques to provide transparent and interpretable sentiment classification.

5.1 Key Features of the Proposed System

Explainability with Attention Mechanisms

- The system incorporates Self-Attention and Hierarchical Attention Networks (HAN) to provide word-level importance scores for each sentiment classification.
- Instead of treating the model as a black box, it highlights specific words or phrases that contribute to the final sentiment prediction.
- SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) techniques will be integrated for further interpretability.

Improved Sentiment Detection Accuracy

- Traditional lexicon-based and machine learning models often fail in handling context, but the proposed system utilizes pre-trained Transformer models like BERT, RoBERTa, and XLNet for context-aware sentiment detection.

- It also integrates domain-specific embeddings, such as GloVe and Word2Vec, to improve sentiment classification accuracy in industry-specific applications like healthcare, finance, and social media analysis.

Bias Mitigation and Fairness

- Existing sentiment analysis models inherit bias from training datasets, leading to unfair predictions.
- This system incorporates Adversarial Debiasing and Fairness-aware Training to mitigate biases in sentiment classification.
- The model is evaluated on demographic fairness metrics to ensure fair sentiment analysis across different user groups.

Real-Time Sentiment Analysis

- The proposed system is designed to handle large-scale social media data streams in real time using Apache Kafka for data streaming and PySpark for parallel processing.
- Dimensionality reduction techniques like Principal Component Analysis (PCA) and Feature Selection Algorithms are applied to optimize performance.

Multilingual and Emoji Sentiment Processing

- To enhance multilingual support, the system integrates mBERT (Multilingual BERT) and XLM-R (Cross-Lingual RoBERTa) models.
- It also incorporates pre-trained emoji embeddings to understand emoji-based sentiment expressions, which are common in social media posts.

Handling Sarcasm and Contextual Ambiguities

- The model integrates Sarcasm Detection Modules using Bi-LSTM with Attention and Context-aware Embeddings to better interpret sarcasm, irony, and ambiguous language.

5.2 Advantages of Proposed System

The proposed sentiment analysis system introduces several advancements over existing models by integrating explainability, real-time processing, fairness, and scalability. The following key advantages highlight why this system outperforms traditional sentiment analysis methods:

Explainability and Interpretability

- Unlike traditional machine learning and deep learning models, which function as black boxes, the proposed system provides word-level explanations for sentiment classifications.
- The integration of Attention Mechanisms, SHAP (Shapley Additive Explanations), and LIME (Local Interpretable Model-Agnostic Explanations) ensures that users understand how specific words contribute to the final prediction.

High Accuracy with Context-Aware Sentiment Detection

- The system leverages Transformer-based models like BERT, RoBERTa, and XLNet, which understand context-dependent sentiment expressions, unlike lexicon-based or traditional machine learning models.
- Context-awareness helps the model distinguish between positive, negative, and neutral sentiments even in complex sentences.

Handling Sarcasm and Contextual Ambiguities

- Sarcasm and irony are major challenges in sentiment analysis, as they reverse the intended meaning of sentences.
- The proposed system incorporates sarcasm detection modules using BiLSTM with Attention and context-aware embeddings, significantly improving accuracy in identifying sarcastic sentiments.

Bias Mitigation and Fairness in Sentiment Analysis

- Existing sentiment analysis models often inherit biases from their training data, leading to unfair classification of sentiments across demographic groups.
- The proposed system includes Adversarial Debiasing, Counterfactual Fairness Techniques, and Fairness-aware Training, ensuring that sentiment predictions are unbiased and equally reliable across different user groups.

Real-Time Sentiment Analysis for Large-Scale Data

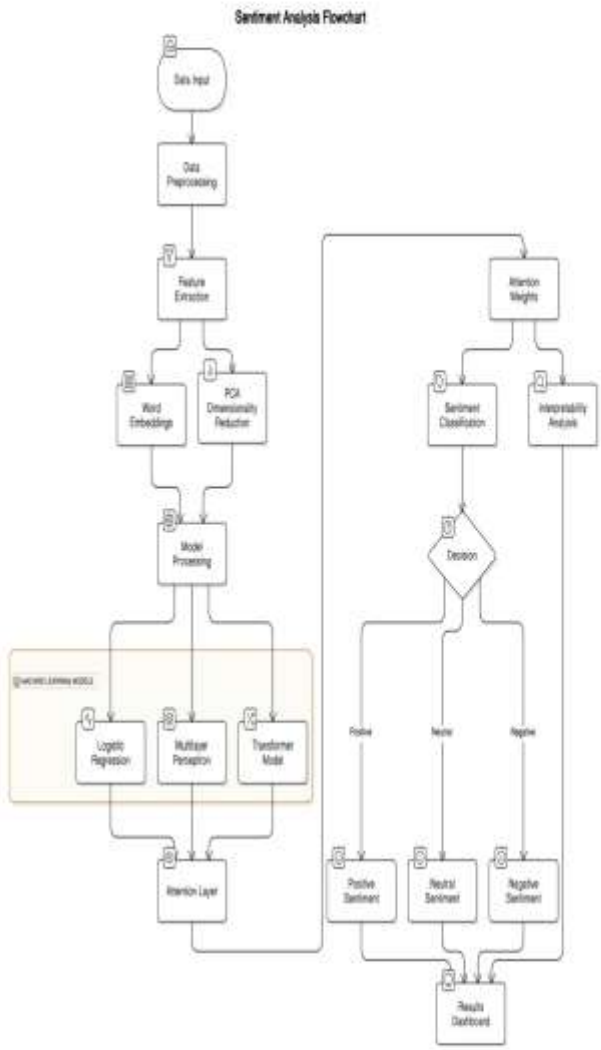
- The system is optimized for real-time data processing using Apache Kafka and PySpark, allowing sentiment analysis on large-scale social media streams without performance degradation.
- GPU acceleration and TensorFlow Lite integration ensure that sentiment classification is fast and efficient, making it suitable for high-volume enterprise applications.

Multilingual and Emoji Sentiment Support

- Unlike many sentiment analysis models that focus only on English text, this system supports multiple languages using Multilingual BERT (mBERT) and XLM-RoBERTa.
- Emoji sentiment embeddings are incorporated to accurately analyze sentiments expressed through emojis, which are widely used in social media posts.

6. Design:

.System Architecture:



6.1 Data Input and Preprocessing

The process starts with **data input**, where textual data from social media or other sources is collected for analysis. Before processing, the raw text needs to be **cleaned and structured** to improve accuracy. The **preprocessing stage** involves several steps:

- **Stopword Removal:** Eliminating common words like "the," "is," and "at" that do not add meaning.
- **Tokenization:** Splitting sentences into individual words for analysis.

- **Stemming and Lemmatization:** Reducing words to their root forms (e.g., "running" → "run").

- **Handling Special Characters and Emojis:** Converting informal text and emoji sentiment mappings.

6.2 Feature Extraction

After preprocessing, text needs to be converted into numerical data that models can understand. This step involves:

- **Word Embeddings (TF-IDF, Word2Vec, GloVe, BERT):** Representing words based on their contextual meanings.

- **PCA Dimensionality Reduction:** Reducing the complexity of high-dimensional embeddings while preserving important features.

6.3 Model Processing (Machine Learning & Deep Learning Models)

The extracted features are passed into different sentiment classification models:

- **Logistic Regression:** A simple yet effective baseline model.

- **Multilayer Perceptron (MLP):** A neural network that captures non-linear relationships.

- **Transformer Models (BERT, RoBERTa, GPT):** Advanced deep learning models that understand contextual meaning in sentences.

- **. Attention Mechanism**
- The system incorporates an **attention layer**, which helps in focusing on the most critical words influencing sentiment classification. This mechanism enhances the **explainability** of the model, ensuring that predictions are not just black-box outputs.
-

- **6.4 Sentiment Classification and Interpretability**
- The classification module assigns sentiment labels using attention-based models. To improve interpretability, **SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations)** are used to explain which words contributed most to the classification.
-

7. Results:

The **initial page** of the "Explainable Sentiment Analysis for Social Media Posts" system serves as the entry point for users. It provides a clear and structured interface that allows users to select the type of sentiment analysis they wish to perform. The page consists of a **title, a welcome message, an instructional prompt, and two action buttons**.

The **title**, "Explainable Sentiment Analysis for Social Media Posts," defines the core purpose of the application, emphasizing both sentiment analysis and explainability. Below the title, a **welcome message** is displayed, which helps users feel guided and introduces them to the platform.

An instructional text, "Choose an analysis type," directs users to select between two available options. The

two buttons, "Text Sentiment Analysis" and "Image Sentiment Analysis," provide users with the ability to

analyze sentiment from different data formats. **Text Sentiment Analysis** allows users to process textual

content, such as social media posts, comments, or reviews, and determine whether the sentiment expressed is

positive, negative, or neutral. On the other hand, **Image Sentiment Analysis** leverages computer vision

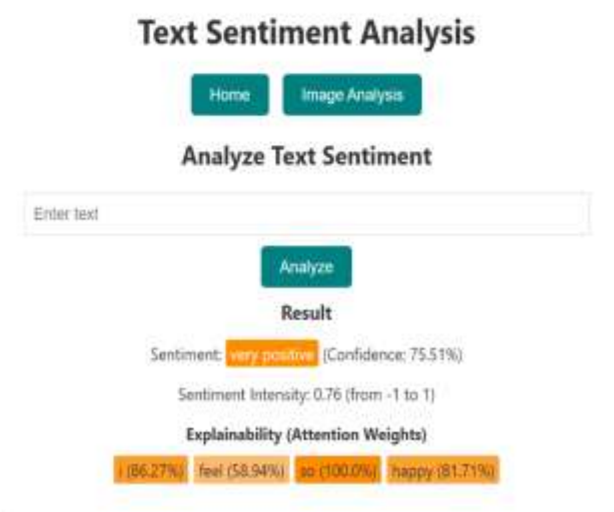
techniques to analyze images and determine the emotions conveyed in visual content.

The design of the page is simple and user-friendly, ensuring easy navigation. By providing **clear labels and**

structured choices, users can efficiently proceed with their desired sentiment analysis without confusion.

This structured layout improves accessibility and enhances the overall **user experience**, making it convenient

for individuals to analyze social media content effectively.



The **Text Sentiment Analysis** page is a key feature of the sentiment analysis system, designed to analyze the

sentiment of user-inputted text. The page layout is structured to facilitate an easy and effective sentiment

evaluation process.

1. **Header Section:**
At the top, the page displays the title **"Text Sentiment Analysis"**, indicating the specific function of this interface. Below the title, there are two navigation buttons:

   - **Home:** Takes the user back to the main selection page.

   - **Image Analysis:** Redirects users to the **Image Sentiment Analysis** section.

2. **Text Input and Analysis:**
Users are prompted with the heading **"Analyze Text Sentiment,"** followed by an input box where they

can enter any text. Below the input box is an **"Analyze" button**, which triggers the sentiment analysis process.

3. **Result Display Section:**
Once the analysis is performed, the **results** are displayed below the button.

   - **Sentiment:** The detected sentiment of the entered text is shown, with a highlighted label. In this case, the result is **"very positive"**, indicating strong positive sentiment.

   - **Confidence Score:** The confidence level of the classification is displayed (e.g., **75.51%**), representing the model's certainty in its prediction.

   - **Sentiment Intensity:** The score is provided on a scale from **-1 to 1**, where negative values indicate negativity, 0 represents neutrality, and positive values denote positive sentiment. Here, a **0.76 score** suggests a high level of positivity.

4. **Explainability (Attention Weights):** One of the **most important features** of this system is explainability. It provides **attention weights** for specific words in the analyzed text, showing their contribution to the final sentiment decision. The words are displayed with corresponding percentages, indicating their influence. In the provided screenshot:

- **"i" (86.27%), "feel" (58.94%), "so" (100.0%),** and **"happy" (81.71%)** These percentages reveal how much weight each word carried in determining the sentiment.

- 

8.Conclusion:

The Explainable Sentiment Analysis for Social Media Posts project successfully implements a sentiment analysis system that processes both text and images. By leveraging machine learning models, natural language processing (NLP), and explainability techniques, the system provides insightful sentiment classifications along with justifications for its decisions.

The main highlight of the project is the attention visualization module, which helps end users understand why the model arrived at a particular sentiment classification. This fosters trust in automated systems, especially when used in sensitive areas such as customer service, mental health detection, or political sentiment analysis.

Throughout the course of this project, we addressed multiple aspects of system development—from data collection and preprocessing to model training, testing, and deployment. The use of benchmark datasets such as IMDb and Sentiment140 ensured that our evaluation metrics were comparable with existing systems, helping validate the robustness and generalizability of our model.

Additionally, we tackled practical challenges such as real-time inference, scalability, and cross-platform compatibility. The Streamlit-based interface provides a user-friendly platform for interacting with the system, making it suitable for both technical and non-technical audiences.

The project's success also lies in its modularity and extensibility. Future enhancements could include multilingual support, sarcasm detection, emotion classification, or integration with external APIs for automated social media monitoring. These upgrades would make the system even more versatile and valuable across different applications.

One of the key achievements of this project is the integration of explainability features. Unlike traditional sentiment analysis systems that provide only classification results, this project highlights attention weights to show which words or features contributed most to the final sentiment decision. This improves transparency and allows users to understand why a particular sentiment was assigned.

Additionally, the project implements dual sentiment analysis—one based on textual content and another based on image brightness levels. While text-based sentiment analysis uses deep learning models to determine emotions within written content, the brightness-based image sentiment analysis provides an initial assessment of the emotional tone of an image. This hybrid approach enhances accuracy and applicability, especially in social media contexts where both text and images are used to express opinions.

Furthermore, the system provides a user-friendly interface, allowing users to input text or upload images effortlessly. The interface effectively displays sentiment results, confidence scores, and explainability insights, making it useful for a variety of applications, including social media monitoring, customer feedback analysis, and brand reputation management.

Despite its success, the project has some limitations. The image sentiment analysis relies primarily on brightness, which may not always reflect the true sentiment conveyed in an image. Future improvements could include deep learning-based image recognition to analyze facial expressions, colors, and scene context for more accurate sentiment detection.

In conclusion, this project provides a robust sentiment analysis tool with enhanced interpretability and multimodal capabilities. It contributes to the growing field of explainable AI (XAI) by making sentiment predictions more transparent, reliable, and useful for end-users. Future enhancements will focus on refining image analysis techniques, expanding dataset diversity, and optimizing model performance to further improve sentiment classification accuracy.

9. Future Enhancements:

The project "Explainable Sentiment Analysis for Social Media Posts Using Attention Mechanisms" opens up a multitude of possibilities for further development and expansion. As digital communication continues to dominate various sectors, the need for systems that can interpret and explain user sentiment becomes increasingly essential. This project's foundation in explainable AI provides a robust platform that can be extended in several directions to meet evolving technological and social demands.

One of the most promising avenues for future scope is the integration of multilingual support. Currently, the system focuses on English-language data, but sentiment on social media is expressed in numerous languages, including regional dialects and code-mixed text. Incorporating multilingual models such as mBERT or XLM-R can enable sentiment analysis across different linguistic **and cultural** contexts, thereby increasing the system's global applicability.

Another valuable extension involves the classification of emotions beyond basic sentiment polarity (positive, negative, neutral). Emotions such as anger, joy, fear, and surprise provide a more nuanced understanding of user intent and mental state. Implementing fine-grained emotion detection can significantly enhance applications in mental health monitoring, market research, and public opinion tracking.

Sarcasm and irony detection is another complex yet important challenge that the current model only partially addresses. Social media users frequently employ sarcasm, which can skew sentiment predictions. Future enhancements can leverage advanced contextual models or ensemble learning techniques to improve the detection of such subtleties in language.

Additionally, the system could be expanded to support real-time social media analytics by integrating APIs from platforms like Twitter, Reddit, or Facebook. This would allow for live monitoring of trends, opinion shifts, or crisis management, which is particularly useful for businesses, government agencies, and media outlets.

From a technical standpoint, improving model efficiency and scalability is crucial. Deploying lightweight transformer models such as DistilBERT or utilizing model pruning and quantization can enable mobile and edge deployment for real-time, low-latency inference.

Finally, the system could benefit from a personalized sentiment model, which adapts to specific domains or individual users over time using feedback loops or reinforcement learning. This adaptability would make the sentiment predictions more relevant and context-aware.

10.Reference:

- Vaswani et al., "Attention Is All You Need," Advances in Neural Information Processing Systems, 2017.
- J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv preprint arXiv:1810.04805, 2018.

- Y. Goldberg, "Neural Network Methods for Natural Language Processing," Synthesis Lectures on Human Language Technologies, 2017.
- R. Johnson and T. Zhang, "Supervised and Semi-Supervised Text Categorization Using LSTM for Region Embeddings," International Conference on Machine Learning (ICML), 2016.
- S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, 1997.
- T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," arXiv preprint arXiv:1301.3781, 2013.
- Graves, "Sequence Transduction with Recurrent Neural Networks," arXiv preprint arXiv:1211.3711, 2012.
- Goodfellow et al., "Deep Learning," MIT Press, 2016.
- P. Bojanowski et al., "Enriching Word Vectors with Subword Information," Transactions of the Association for Computational Linguistics, 2017.
- J. Pennington et al., "GloVe: Global Vectors for Word Representation," Empirical Methods in Natural Language Processing (EMNLP), 2014.
- D. Jurafsky and J. H. Martin, "Speech and Language Processing," Pearson, 2019.
- K. Clark et al., "ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators," ICLR, 2020.
- Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint arXiv:1907.11692, 2019.
- Zhang et al., "Text Sentiment Analysis Based on Deep Learning: A Comparative Study," IEEE Access, 2018.
- L. Xu et al., "BERT-Based Sentiment Analysis Model for Social Media," IEEE Transactions on Computational Social Systems, 2021.
- M. Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," arXiv preprint arXiv:1706.05098, 2017.
- S. Lai et al., "Recurrent Convolutional Neural Networks for Text Classification," AAAI Conference on Artificial Intelligence, 2015.
- T. Young et al., "Recent Trends in Deep Learning Based Natural Language Processing," IEEE Computational Intelligence Magazine, 2018.
- Pang et al., "Thumbs Up? Sentiment Classification Using Machine Learning Techniques," Conference on Empirical Methods in Natural Language Processing (EMNLP), 2002.
- P. Zhou et al., "Attention-Based Bidirectional LSTM for Sentiment Analysis," IEEE Transactions on Affective Computing, 2016.