WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

LLM-Enhanced Emotional Intelligence for Text Processing

Lenka Swathi[1], Menda Usha[2], Mygapu Kamal Tej[3], Reshma Shaik[4], Sappa Prakash[5]

[1]Lenka Swathi, Information Technology, Rajam, Andhra Pradesh
[2]Menda Usha, Information Technology, Rajam, Andhra Pradesh
[3]Mygapu Kamal Tej, Information Technology, Rajam, Andhra Pradesh
[4]Reshma Shaik, Information Technology, Rajam, Andhra Pradesh
[5]Sappa Prakash, Information Technology, Rajam, Andhra Pradesh

ABSTRACT :

This application harnesses the power of large language models to efficiently extract key information and emotional insights from lengthy documents. Developed using Streamlit, it leverages advanced natural language processing techniques to perform content summarization and emotion analysis across various input formats, including PDFs, text files, and images. By utilizing the BART model, long documents are intelligently segmented to generate coherent and high-quality summaries. Simultaneously, a pre-trained transformer model assesses the emotional tone of the content. Integrated optical character recognition (OCR) allows for seamless text extraction from images, enhancing the tool's versatility. Designed with a user-friendly interface and real-time processing capabilities, this solution significantly reduces manual effort, making it highly beneficial for students, researchers, and professionals. To enhance flexibility, Optical Character Recognition (OCR) powered by Tesseract is integrated, enabling text extraction from scanned documents or image files (JPEG, PNG, etc.). This ensures that non-digital or handwritten sources can also be analyzed, broadening the tool's scope. Future upgrades aim to incorporate multilingual support, refine OCR performance, and enable cloud-based integration—further enhancing accuracy, scalability, and efficiency.

Key Words: Text Summarization, Sentiment Analysis, Streamlit, BART Model, Transformer Models, OCR, PDF Processing, Real-time Processing

INTRODUCTION

In today's digital age, vast amounts of unstructured textual data are generated across platforms such as academic publications, business reports, emails, social media, and scanned documents. Processing and interpreting this data manually is both time-consuming and inefficient. As a result, there is a growing demand for intelligent tools that can automatically extract key insights, generate concise summaries, and evaluate the emotional tone embedded within the content. Recent breakthroughs in artificial intelligence, particularly with large language models (LLMs), have revolutionized the field of natural language processing (NLP), enabling machines to understand and interpret human language with unprecedented accuracy. This project presents a robust and interactive application, developed using the Streamlit framework, that harnesses the power of advanced NLP techniques to perform automated text summarization and sentiment analysis across diverse input formats—including PDF documents, plain text files, and image-based content. The application is powered by the BART (Bidirectional and Auto-Regressive Transformers) model, a state-of-the-art encoder-decoder architecture capable of generating high-quality summaries from long-form texts. To assess sentiment, it integrates a fine-tuned transformer model (such as RoBERTa or DistilBERT), which classifies content into emotional categories like positive, negative, neutral, or more nuanced feelings such as joy, anger, or sadness.

To accommodate image-based inputs, the application incorporates Optical Character Recognition (OCR) using Tesseract, enabling the extraction of text from scanned documents, photos, or handwritten notes. This multi-format compatibility ensures that users can process a wide range of documents seamlessly, without being limited by input type.The tool is designed for a broad audience students looking to summarize research papers, professionals analyzing customer feedback or reports, and researchers studying sentiment trends across textual datasets. Its user-friendly interface, real-time response, and modular architecture make it both accessible and scalable. For sentiment and emotion analysis, the system utilizes transformer-based models such as RoBERTa**,** DistilBERT, or FinBERT, depending on the context and domain, providing sentiment classification across predefined emotional states (e.g., positive, negative, neutral) or more granular categories such as joy, fear, surprise, anger, and sadness. These models are pre-trained on massive corpora and fine-tuned on emotion-labeled datasets like GoEmotions or EmotionX.

To expand its capabilities further, the application incorporates Optical Character Recognition (OCR) using the Tesseract engine, allowing it to convert non-digital or scanned images into machine-readable text. This makes it possible to analyze documents captured by mobile devices, handwritten notes, and screenshots—a critical feature for educational, archival, or field research settings.

LITERATURE SURVEY

This research aims to improve abstractive text summarization in the Urdu language by utilizing a comprehensive dataset comprising 19,615 document-summary pairs. The dataset was developed by translating an open-source English corpus into Urdu using Google Translate, followed by thorough manual refinement to ensure linguistic precision and natural fluency. To generate contextually rich summaries, the study leverages transformer-based models capable of capturing dependencies across multiple sentences, resulting in more coherent and meaningful outputs. Furthermore, a new evaluation metric the Context-Aware RoBERTa Score is proposed. This metric evaluates summaries based on contextual relevance and logical flow, offering a more robust alternative to traditional scoring techniques that often fall short in measuring narrative consistency.[1]

This paper furthermore implements, the implementation of transfer learning greatly enhances the summarization model by leveraging pretrained language models and embeddings, leading to more accurate and context-aware outputs. The system combines multiple abstractive summarization strategies—including recurrent neural networks (RNNs), attention-based methods, and advanced transformer models like PEGASUS to produce high-quality summaries.
In addition to summarization, this study explores sentiment classification by categorizing textual content into positive, negative, and neutral sentiments. A BERT-based model is fine-tuned specifically for analyzing sentiment and emotional tone within Hungarian political speeches. The effectiveness of this sentiment analysis is assessed using standard evaluation metrics, including Precision, Recall, and F1-Score, ensuring a comprehensive understanding of model performance.[2]

In this paper, emotion classification is applied to categorize emotions such as fear, anger, joy, and sadness within textual data. To enhance the model's accuracy, confusion matrices are used to identify where misclassifications occur, which aids in improving the system's overall reliability. By incorporating pretrained universal embeddings, the model is able to grasp subtle nuances in context, leading to better sentiment analysis results in political text. Advanced deep learning techniques further fine-tune the evaluation of emotional content in political speeches and statements.
The study also investigates Arabic sentiment analysis, including the detection of sarcasm and dialect classification, employing a variety of machine learning algorithms. The classifiers tested include K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, and Naïve Bayes, with a comprehensive evaluation of their strengths and weaknesses in handling these complex challenges.[3]

VARIOUS METHODS ARE USED IN LLM-ENHANCED EMOTIONAL INTELLIGENCE FOR TEXT PROCESSING:

A. Bidirectional Encoder Representations from Transformers(BERT)

The BERT (Bidirectional Encoder Representations from Transformers) model is a pre-trained transformer-based architecture designed to understand the context of words in a sentence by considering both the left and right context. Unlike traditional models that process text in a unidirectional manner, BERT reads text bidirectionally, enhancing its ability to capture the nuances of language. It is pre-trained on vast amounts of text data using two main tasks: Masked Language Modeling (MLM) and Next Sentence Prediction (NSP). MLM involves randomly masking words in a sentence and training the model to predict them, while NSP helps BERT understand relationships between sentences. BERT has been fine-tuned for various NLP tasks like sentiment analysis, question answering, and text classification. Its deep contextualized word embeddings make it highly effective for understanding complex linguistic patterns. By leveraging pre-trained knowledge, BERT can be adapted to a wide range of downstream tasks with minimal task-specific data. The model is efficient, scalable, and has set new benchmarks in multiple NLP tasks, including question answering and language inference.

B. Natural Language Processing (NLP)

Natural Language Processing (NLP) is essential for extracting meaningful insights from patient inquiries, enhancing the quality of responses provided by chatbots. The NLP process involves several key stages. Initially, text preprocessing takes place, where the input text is tokenized, irrelevant stopwords are removed, and the text is normalized to ensure consistency across data. Following this, Named Entity Recognition (NER) identifies important medical terms, symptoms, and diagnoses within the patient's query. The chatbot then proceeds with intent detection, categorizing the query into relevant topics such as diagnosis, treatment, or appointment scheduling. Next, sentiment analysis is conducted to assess the emotional tone of the query, helping the system understand the urgency of the patient's concern. Finally, the chatbot generates a response by leveraging pre-trained language models such as GPT or BERT, ensuring that the reply is contextually appropriate and meaningful. While NLP greatly improves chatbot performance and interaction with patients, it requires access to large, annotated datasets to accurately interpret and process medical-related text for effective communication.

C. Large Language Models (LLM)

Large Language Models **(**LLMs**)** are sophisticated deep learning systems designed to understand, interpret, and generate human language by training on large and varied text datasets. The creation of LLMs begins with the collection and preprocessing of vast amounts of text, which involves cleaning and tokenizing the data into smaller units such as words or subwords for efficient processing. After this, a deep learning architecture—typically the Transformer model—is chosen, and key hyperparameters, such as the number of layers and attention heads, are set. The model is then trained using optimization algorithms like Adam or Stochastic Gradient Descent (SGD), often requiring significant computational resources due to the volume of data and the complexity of the model.

Once training is complete, the model's performance is assessed using metrics such as perplexity and accuracy on a separate test dataset to gauge its effectiveness. To specialize the model for specific tasks or industries, it is further fine-tuned using smaller, domain-specific datasets. After fine-tuning, the LLM is ready for deployment in practical applications such as text generation, summarization, and conversational AI. Following deployment, ongoing performance monitoring, user feedback collection, and periodic retraining with updated data are essential for maintaining and enhancing the model's capabilities over time. Although LLMs are powerful tools for a wide range of natural language processing tasks, they require considerable computational power and fine-tuning to ensure they deliver optimal performance.

RESULTS:

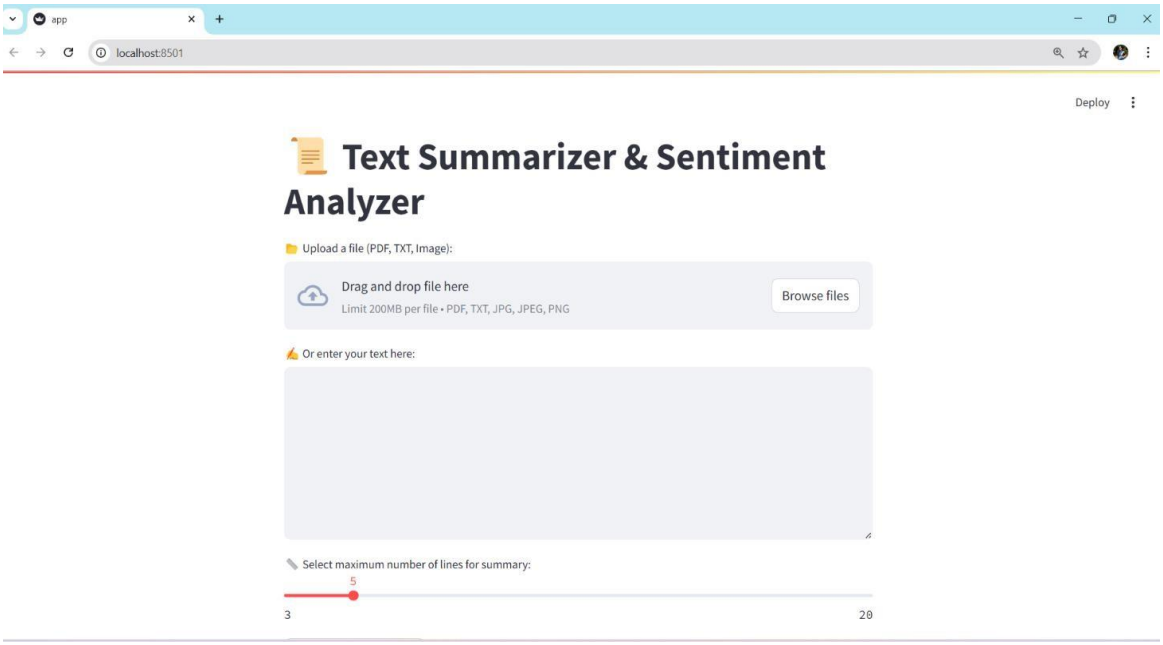

Figure 4.5: Streamlit interface

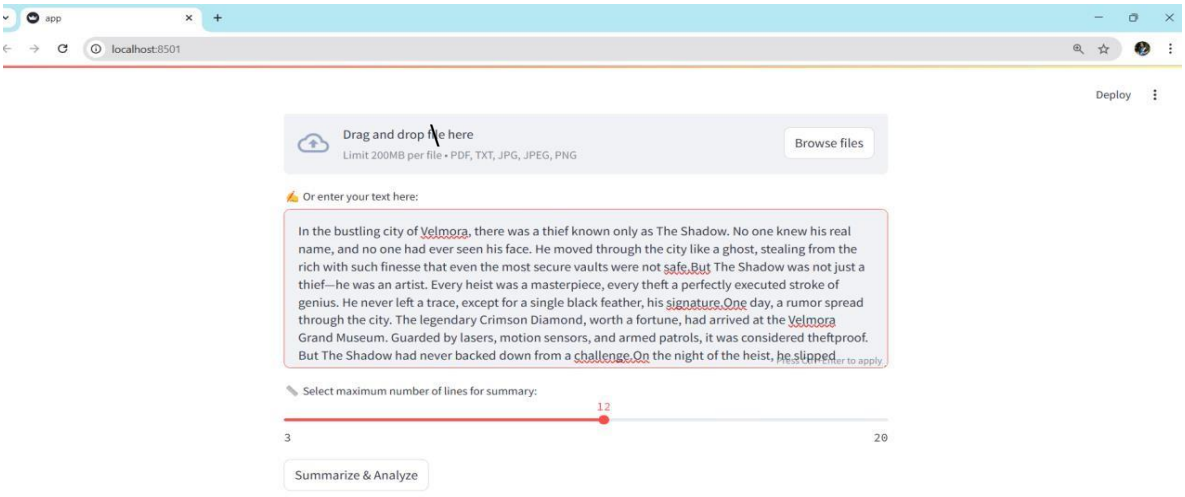

Figure 4.6: Input

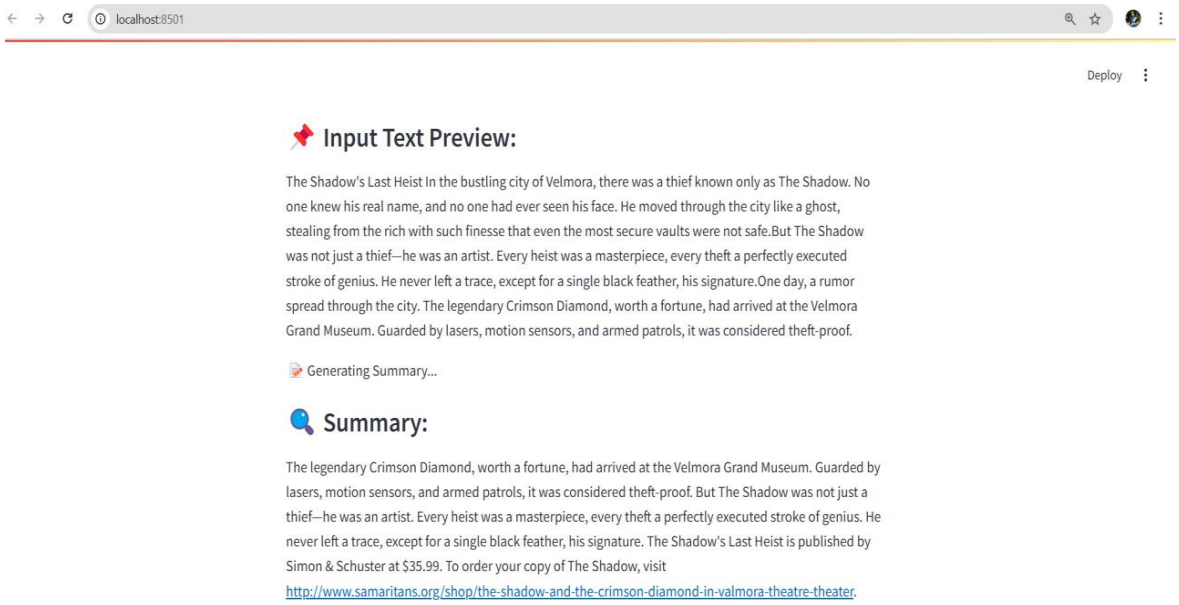

Figure 4.7: Output

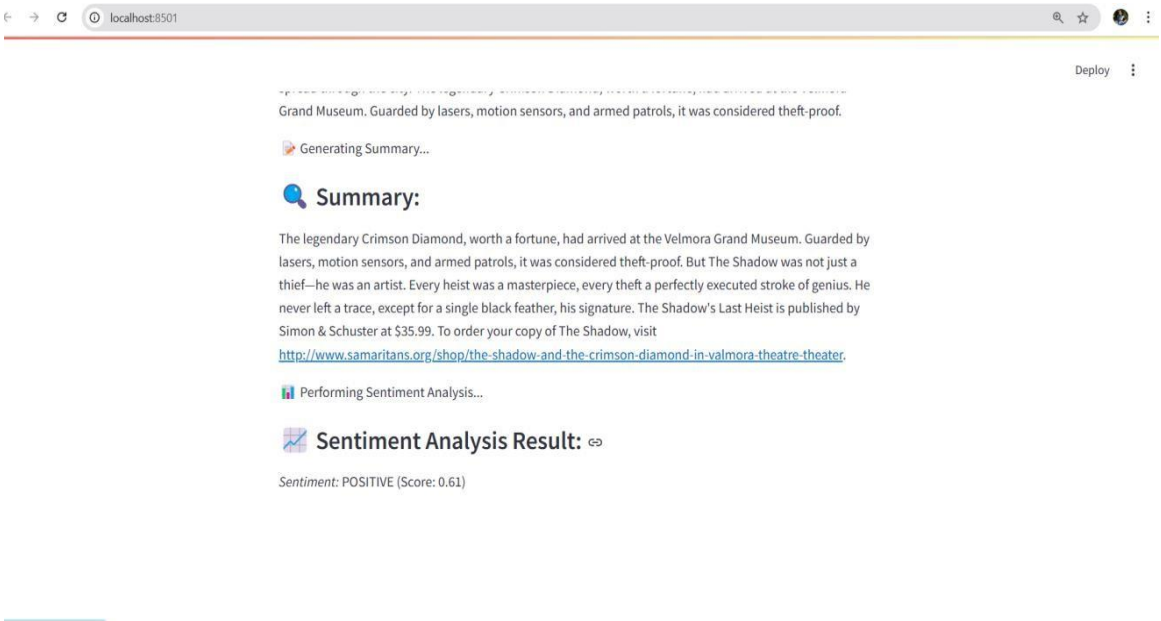

Figure 4.8: Results

CONCLUSION

This study Streamlit application offers a robust and efficient solution for both text summarization and sentiment analysis, supporting multiple file formats including PDFs, text files (.txt), and images. By utilizing the BART model a powerful transformer-based model known for its exceptional performance in abstractive summarization the application generates clear and coherent summaries of long documents. Additionally, the system integrates sentiment analysis pipelines to provide valuable insights into the emotional tone of the content, categorizing text into positive, negative, or neutral sentiments. To enhance the overall user experience, the application is equipped with comprehensive error-handling mechanisms that address common issues such as unsupported file types, extraction failures, or corrupted text. This ensures smooth interaction with the tool, even in cases of unexpected input scenarios. The system's text chunking capabilities allow it to process large documents efficiently, breaking them into manageable sections that maintain contextual integrity for both summarization and sentiment analysis. Furthermore, model caching is incorporated to optimize processing times, ensuring that frequently accessed data or models are reused without redundant computations, ultimately improving responsiveness, particularly when working with larger files. In summary, this versatile application combines advanced natural language processing techniques with an intuitive interface, making it a powerful tool for users seeking to extract concise summaries and perform sentiment analysis across a wide variety of text sources. Its scalability, performance optimizations, and error resilience position it as a valuable resource for researchers, students, and professionals alike.

REFERENCES:

[1] Pilault, J., Li, R., Subramanian, S., & Pal, C. (2020, November). On extractive and abstractive neural document summarization with transformer language models. In Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) (pp. 9308-9319).

[2] Wahab, M. H. H., Ali, N. H., Hamid, N. A. W. A., Subramaniam, S. K., Latip, R., & Othman, M. (2023). A review on optimization-based automatic text summarization approach. IEEE Access, 12, 4892-4909.

[3] Kotkar, A. D., Mahadik, R. S., More, P. G., & Thorat, S. A. (2024, August). Comparative Analysis of Transformer-based Large Language Models (LLMs) for Text Summarization. In 2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET) IEEE.

[4] Işıkdemir, Y. E. (2024). NLP TRANSFORMERS: ANALYSIS OF LLMS AND TRADITIONAL APPROACHES FOR ENHANCED TEXT SUMMARIZATION. Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi, 32(1), 1140-1151.

[5] Zhang, H., Yu, P. S., & Zhang, J. (2024). A Systematic Survey of Text Summarization: From Statistical Methods to Large Language Models. arXiv preprint arXiv:2406.11289.