International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-779

HYBRID DEEP LEARNING FRAMEWORK FOR EMOTION DETECTION IN SOCIAL MEDIA POSTS

Dhivya Prasath K, Kowsalya S, Malini V, Aswin S

Dhivya Prasath K, Dept. of Computer Science and Engineering, Anna University, IN
Kowsalya S, Dept. of Computer Science and Engineering, Anna University, IN
Malini V, Dept. of Computer Science and Engineering, Anna University, IN
Aswin S, Dept. of Computer Science and Engineering, Anna University, IN

Abstract - *This project develops a hybrid deep learning framework for emotion detection in social media posts, combining CNNS and LSTMS to enhance accuracy by capturing local and sequential features in text data. Pre-trained embeddings like Word2Vec and GloVe are used for better semantic understanding, while attention mechanisms and Transformers improve model performance. The framework integrates facial expression recognition through CNNS, creating a multi-modal approach for robust emotion detection. Implemented in Python using TensorFlow and Keras, the model is evaluated on diverse, multi-lingual datasets collected via APIs and web scraping tools from various social media platforms. Standard metrics such as accuracy, precision, recall, and Fl-score are used for evaluation. The project aims to address challenges like slang and informal language, ensuring the model is adaptable and effective across diverse social media environments.*

Key Words: H*ybrid deep learning, Emotion detection, Social media posts, Attention mechanisms, Transformers, Multi-modal approach, LSTM, Word2Vec, GloVe, CNN.*

1. INTRODUCTION

Emotion detection in social media has become an essential area of research due to the increasing volume of online interactions. Accurately identifying emotions in posts can provide valuable insights for sentiment analysis, mental health monitoring, and customer feedback analysis. However, the informal nature of social media, characterized by slang, diverse expression styles, and multilingual content, poses challenges for traditional approaches. This project proposes a hybrid deep learning framework combining Convolutional Neural Networks (CNNS) and Long Short-Term Memory (LSTM) networks to effectively detect emotions by leveraging both textual and facial cues.

The framework uses CNNS to capture local patterns in facial expressions and text, while LSTMS are employed to understand sequential dependencies in the text, ensuring a comprehensive analysis. Pre-trained embeddings such as Word2Vec and GloVe are integrated to improve the semantic understanding of informal language. Attention mechanisms and Transformers are also explored to enhance the model's interpretability and focus. Implemented using Python, TensorFlow, and Keras, the framework is tested on diverse datasets obtained from social media platforms, ensuring robustness across different languages and text styles. By addressing the complexities of social media communication, this project aims to create an accurate and efficient emotion detection system suitable for real-world applications.

1.1 Background of the Work

The rapid growth of social media has made it a primary platform for communication, creating a need for accurate emotion detection to understand user sentiments. Existing solutions often focus on either text or facial data, limiting their ability to provide comprehensive insights. Social media data presents unique challenges, including slang, informal language, and multilingual content, which traditional models struggle to process effectively. Current systems lack the adaptability required for diverse and evolving online interactions. This invention addresses these gaps by introducing a hybrid framework that integrates text and facial cues. By leveraging deep learning techniques such as CNNs and LSTMs, along with pre-trained embeddings and attention mechanisms, the framework enhances semantic understanding and interpretability. This innovation ensures robust emotion detection across varied social media contexts, filling a critical gap in the field

1.2 Motivation and Scope of the Proposed Work

Hybrid Architecture: Combines CNNS for facial expression recognition with LSTMS for sequential text analysis, improving overall emotion detection accuracy. Multi-Modal Input: Integrates both textual and facial cues to provide a comprehensive understanding of user emotions, overcoming the limitations of single-modality approaches. Advanced Pre-trained Embeddings: Utilizes embeddings like Word2Vec and GloVe to enhance the

International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-779

model's semantic understanding of informal language and slang prevalent in social media. D O Attention Mechanisms: Incorporates attention layers to allow the model to focus on critical features in the data, enhancing interpretability and performance during emotion detection. O Multi-Lingual Robustness: Designed to effectively handle diverse, multi-lingual datasets, ensuring adaptability and accuracy in emotion detection across various social media platforms.

2. METHODOLOGY

The methodology for this project employs a hybrid deep learning framework combining CNNs and LSTMs to capture local and sequential text features. Pre-trained embeddings like Word2Vec and GloVe, along with attention mechanisms and Transformers, enhance semantic understanding. A multi-modal approach integrates text and facial expression recognition, evaluated on diverse, multi-lingual datasets using standard metrics.

2.1 System Architecture

The system architecture consists of a text processing module leveraging pre-trained embeddings (Word2Vec, GloVe) for semantic representation and a deep learning layer combining CNNs and LSTMs to extract local and sequential features. An attention mechanism and Transformer layer enhance context understanding. The facial expression recognition module, powered by CNNs, integrates with the text module for a multi-modal approach. The combined outputs are fed into a classification layer to detect emotions, evaluated using accuracy, precision, recall, and F1-score as shown in Fig-1.

2.2 Data Acquisition

Gather raw social media posts using APls and web scraping tools, followed by cleaning, normalization, and tokenization to prepare the data for analysis. Convert cleaned text into numerical vectors using techniques like Bag-of-Words or TF-IDE, and integrate pre-trained embeddings such as Word2Vec and GloVe to capture semantic meanings.

2.3 Model Development

Design and configure a hybrid architecture combining CNNS for feature extraction from facial expressions and LSTMS for analyzing sequential dependencies in text data. Train the hybrid model using labeled datasets, optimize it with algorithms like stochastic gradient descent, and perform hyperparameter tuning to improve performance. Use validation data sets to assess model accuracy and prevent overfitting.

2.4 Performance Evaluation

Evaluate the model's effectiveness using standard metrics such as accuracy, precision, recall, and Fl-score, and ensure robustness across diverse, multi-lingual datasets from various social media platforms.

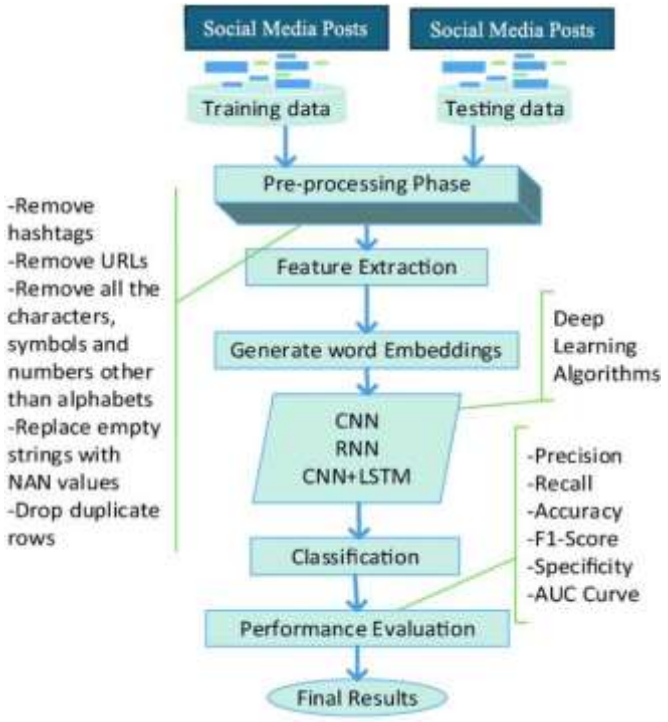

Fig -1- Flowchart

BRIEF DESCRIPTION OF FLOWCHART:

- Input Data: Social media posts are split into training and testing datasets for model building and evaluation.
- Pre-Processing: Text data is cleaned by removing handles, hashtags, URLs, nonalphabetic characters, and duplicates, followed by tokenization.
- Feature Extraction: Key features are extracted from the processed text for further analysis by deep learning models.
- Word Embeddings: Text is converted into dense numerical vectors to represent semantic meaning for model input.
- Deep Learning Models: CNN, RNN, and a hybrid CNN+LSTM models are used for emotion classification.
- Performance Evaluation: Model performance is assessed using metrics like precision, recall, accuracy, F1-score, and AUC.
- Final Results: The framework outputs the detected emotions with insights based on evaluation metrics.

3. CONCLUSIONS

"Hybrid Deep Learning Framework for Emotion Detection in Social Media Posts" seeks to enhance the accuracy and efficiency of emotion recognition by integrating image and text data through advanced deep learning techniques. Social media platforms serve as vast repositories of user-generated content, where individuals express emotions through textual posts, images, and videos. Recognizing these emotions accurately can provide valuable insights for

various applications, such as brand monitoring, mental health analysis, and societal sentiment tracking. In this framework, Convolutional Neural Networks (CNNs) are employed for image analysis, effectively capturing emotional cues from visuals, such as facial expressions, color usage, and contextual imagery. CNNs excel at feature extraction from images, enabling the model to understand complex visual patterns associated with specific emotions. Meanwhile, for text analysis, a Multinomial Naive Bayes (NB) classifier is used to evaluate the emotional tone of textual content,where it analyzes the occurrence of words and phrases to classify emotions based on their frequency and context. The hybrid model merges the outputs of both CNN and Multinomial NB, creating a comprehensive understanding of emotional expression. By integrating these modalities, the framework capitalizes on the strengths of each approach: CNN's ability to analyze images and the NB's efficiency in processing textual information. The findings of this research could pave the way for more effective tools in analyzing public sentiment and understanding user behavior in digital spaces.

Suggestions for Future Work

1. **Real-Time-Emotion-Detection:**
   Develop and deploy the framework for real-time emotion analysis on live social media streams, ensuring low-latency performance and scalability.
2. **Cross-Domain-Adaptability:**
   Extend the framework to other domains, such as healthcare for mental health monitoring or ecommerce for customer sentiment analysis.
3. **Enhanced-Multi-Lingual-Capabilities:**
   Incorporate advanced multi-lingual models like mBERT or XLM-R to improve emotion detection accuracy across more languages and dialects.

REFERENCES

[1] Zhang, Y., Zhao, L., & Chen, J. (2020). A Hybrid Deep Learning Approach for Emotion Recognition from Text in Social Media. IEEE Access, 8, 32905-32916.

[2] Poria, S., Chan, L., & Liu, Z. (2017). A Hybrid Approach for Emotion Recognition in Text and Images from Social Media. IEEE Transactions on Affective Computing, 8(2), 124-137.

[3] Chen, J., Yang, S., & Hu, X. (2019). Emotion Detection in Social Media Using Multi-modal Deep Learning. Journal of Computer Science and Technology, 34(2), 400-414.

[4] Zhou, Z., & Wang, H. (2018). Hybrid Deep Learning for Emotion Recognition in Social Media Text. Neurocomputing, 273, 159-169.

[5] Kharde, V., & Sonawane, S. (2020). Emotion Detection in Social Media Using Machine Learning and Deep Learning Techniques. International Journal of Computer Applications, 975, 8887.