KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)
Terakreditasi Nomor 204/E/KPT/2022 | Vol. 5, No. 2, April (2024), pp. 737-745

Comparative Analysis of Deep Learning Architectures for Emotion Recognition in Text

Gregorius Airlangga
Universitas Katolik Indonesia Atma Jaya, Indonesia
E-mail: gregorius.airlangga@atmajaya.ac.id

Abstract

This study delves into the intricacies of emotion recognition within textual data, presenting a comprehensive analysis of three prominent deep learning models: Long Short-Term Memory networks (LSTMs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs). Employing a 5-fold cross-validation methodology, the research meticulously evaluates each model's performance in accurately classifying a spectrum of emotions, using metrics such as accuracy, precision, recall, and F1 score. Results indicate that LSTMs outperform their counterparts with an accuracy of 93.48%, closely followed by CNNs at 91.78%, while RNNs lag, showcasing the importance of sophisticated architectural features in handling complex emotional nuances. The study not only highlights the strengths and limitations of each model but also sheds light on the significant role of temporal and contextual understanding in emotion recognition tasks. Through this investigation, we provide insights into the evolving landscape of natural language processing and its capability to decode human emotions, proposing directions for future research in enhancing model performance. This work has broader implications for applications in mental health, customer service, and social media analysis, aiming to refine the interaction between humans and machines in understanding and processing emotional content.

Keywords: *We Emotion Recognition, Deep Learning, LSTM (Long Short-Term Memory), CNN (Convolutional Neural Network), RNN (Recurrent Neural Network)*

1. Introduction

Emotion analysis in textual data has garnered immense interest in the field of natural language processing (NLP) and artificial intelligence (AI), aiming to decipher the underlying sentiments and emotions expressed in written communication [1]–[3]. This burgeoning interest is propelled by the escalating volume of text data generated through social media platforms, customer reviews, and online interactions, necessitating advanced computational methods to understand and interpret human emotions accurately [3]–[5]. The ability to automatically analyze emotions in text has significant implications across various domains, including marketing, customer service, psychotherapy, and social media monitoring, where understanding human sentiments plays a pivotal role in decision-making processes [1], [6], [7]. In the realm of machine learning and deep learning, several models have been proposed and employed to tackle the challenges of emotion analysis, each with its unique strengths and limitations [1], [8], [9]. Traditional approaches, like support vector machines (SVM) and Naïve Bayes, laid the groundwork for text classification tasks [10]. However, the advent of deep learning has shifted the landscape, introducing models capable of capturing complex semantic relationships within text [11]. Among these, Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) are at the forefront, acclaimed for their proficiency in handling sequential data and extracting meaningful patterns from large textual datasets [12].

The literature reveals a plethora of studies emphasizing the efficacy of LSTM, CNN, and RNN models in emotion analysis [13]. LSTM networks, with their capability to

ISSN: 2720-992X

remember long-term dependencies, have shown remarkable success in capturing contextual nuances in text, leading to superior performance in sentiment classification tasks [14]. Conversely, CNNs, primarily recognized for image processing, have been adeptly repurposed for text analysis, demonstrating their strength in extracting local and hierarchical features from text data. RNNs, known for their sequential data processing, have also been integral in analyzing time-series data and text, although they sometimes struggle with long-term dependency recognition [15]. Despite the progress, there exists an urgency to develop more robust and accurate models for emotion analysis, as the current state-of-the-art still faces challenges related to context understanding, sarcasm detection, and the subtleties of emotional expression [16]. These challenges are amplified by the diverse and dynamic nature of human language, including variations in slang, dialect, and cultural nuances, which can significantly affect the interpretation of emotions in text [17].

The goal of this research is to provide a comprehensive and comparative analysis of LSTM, CNN, and RNN models in the context of emotion analysis in text data, aiming to uncover the strengths and weaknesses of each model in capturing and interpreting emotional nuances. This study is prompted by a noticeable gap in systematic comparative analyses that consider various deep learning architectures under consistent experimental conditions, particularly in the realm of emotion analysis. Our contribution to the field is multifaceted. Firstly, we implement a rigorous cross-validation methodology to evaluate each model's performance, ensuring a robust and unbiased assessment. Secondly, we delve into the intricacies of how these models process and analyze emotional content in text, providing insights into their operational mechanisms. Thirdly, we aim to identify specific areas where each model excels or falls short, contributing to a nuanced understanding of their applicability in real-world scenarios. The remainder of this journal article is structured as follows: Section 2 details the literature survey, Section 3 the methodologies employed, including the data preprocessing steps, model configurations, and evaluation metrics. Section 4 presents the results of our experiments, offering a detailed comparative analysis of the performance of LSTM, CNN, and RNN models in emotion analysis. In addition, we also discussed the implications of our findings, providing a critical interpretation of the models' performances and their potential applications in various domains. Finally, Section 5 concludes the article, summarizing the key contributions and insights derived from our research.

2. Literature Survey

Early research in emotion recognition from text relied heavily on lexical databases and rule-based algorithms. Studies like [18] utilized the Linguistic Inquiry and Word Count (LIWC) tool to analyze emotional valence in texts, while [19] leveraged the AFINN lexicon for sentiment analysis. Although these methods provided foundational insights, they were limited by their inability to understand context, the use of fixed lexicons that could not adapt to new expressions or slang, and a general lack of scalability to diverse datasets. The advent of deep learning introduced a paradigm shift in emotion recognition. [20] introduced LSTMs, which significantly improved the model's ability to remember long-term dependencies, a crucial factor in understanding the context in textual data. CNNs, popularized for text analysis by [21] offered another approach by extracting hierarchical features from texts, proving effective in classifying emotions and sentiments. Despite their success, these models often required extensive computational resources and large labeled datasets for training, limiting their accessibility and practicality for smaller projects or less-resourced languages. The introduction of transformer models like BERT [22] and GPT [23] marked another significant milestone, setting new benchmarks in a range of NLP tasks, including emotion recognition. By pre-training on vast corpora, these models achieved remarkable success in capturing the nuances of human language [24]. However, their complexity and the opaque nature of their decision-making processes raised concerns about interpretability and the ethical implications of their use, especially in sensitive applications.

Across the spectrum of studies, several recurring challenges emerge. Firstly, the issue of imbalanced datasets, where certain emotions are underrepresented, remains a significant hurdle, leading to biased models that perform poorly on rare but critical emotional expressions [25]. Secondly, the nuanced nature of human emotions, including the expression of sarcasm, irony, and mixed emotions, poses a persistent challenge to even the most advanced models [26]. Lastly, many studies focus on English language texts, leaving a gap in research applicable to other languages and cultures [27]. Our research is designed to build upon the insights gained from both traditional and deep learning approaches, addressing the noted gaps and challenges. By comparing the effectiveness of LSTMs, CNNs, and RNNs in a structured and comprehensive manner, this study not only elucidates the strengths and weaknesses of each model in recognizing a wide range of emotions but also investigates their performance in a cross-validated setting to enhance model reliability and generalizability. Unique to our approach is the emphasis on cross-validation techniques to assess model performance, providing a robust evaluation framework that is often overlooked in emotion recognition studies. Additionally, by systematically addressing the issue of imbalanced datasets through strategic model training and evaluation, our research contributes to the development of more equitable and effective emotion recognition systems. This is particularly relevant in applications such as mental health monitoring and customer service, where the accurate detection of less frequent emotional states can be critical. This study contributes to the body of knowledge by offering a detailed comparison of model performances across diverse emotions, shedding light on the practical implications of deploying these models in real-world applications. Furthermore, our research extends the dialogue on the ethical and interpretative aspects of using advanced deep learning models for emotion recognition, advocating for a balanced approach that considers accuracy, transparency, and fairness.

3. Research Methodology

The Research Methodology section of a paper on emotion recognition from text using deep learning outlines the systematic approach employed to investigate the research questions. This section is crafted to ensure the study's reproducibility and to provide clarity on the processes used to collect data, preprocess it, select, and implement models, and evaluate their performance. The foundation of this research is built upon a carefully curated dataset, chosen to encompass a broad spectrum of human emotions. This dataset is an amalgamation of texts sourced from diverse public domains, including social media platforms, literary works, and transcriptions of spoken dialogues. The selection of these sources is strategic, aimed at capturing the multifaceted nature of language and how it conveys emotions across different contexts and modalities.

3.1. Dataset Description and Processing

Social media posts, with their candid and spontaneous expressions of feelings, provide a rich vein of real-time emotional content. Literary excerpts, on the other hand, offer a more nuanced and crafted representation of emotions, often portraying complex emotional narratives and interactions. Transcripts of spoken dialogue bring in the dynamics of spoken language, including intonation and speech patterns, which are essential in understanding emotional expressions in conversational contexts. Each piece of text in the dataset has been labeled with one of six primary emotions: joy, sadness, anger, fear, love, and surprise. This labeling process involved multiple annotators, who independently assessed each text snippet to assign the most fitting emotional category. The use of multiple annotators is a deliberate choice to mitigate subjective bias and to ensure a reliable consensus on the emotional labeling, thereby enhancing the dataset's annotation quality. Dataset can be downloaded from [28]. Furthermore, the preprocessing of the collected data is a crucial step, transforming raw text into a format that is amenable to analysis by deep learning models. This preprocessing pipeline is carefully constructed to address the idiosyncrasies of textual data, ensuring that the resulting vectors accurately represent the underlying emotional content. The first step in this pipeline is tokenization, where the text is segmented into individual words or tokens. This process is fundamental to breaking down the complex structure of language into manageable units that can be analyzed and interpreted by the models. Cleaning the data involves removing extraneous elements that do not contribute to or might even obscure the emotional significance of the text. This includes stripping out HTML tags, special characters, and numerical figures, which are typically irrelevant to the task of emotion recognition.

Lowercasing is then applied to standardize the text, ensuring that the same words in different cases are recognized as identical. This uniformity is crucial for maintaining consistency in the dataset and preventing the models from treating capitalized and lowercase versions of the same word as different entities. The removal of stop words is another key preprocessing step. Common words like "the," "is," and "in," which are pervasive in text but generally carry little emotional weight, are filtered out to reduce noise in the data and to allow the models to focus on more meaningful words that are likely to contribute to emotional expression. Lemmatization further refines the dataset by condensing words to their base or dictionary form. Unlike stemming, which merely strips suffixes and can sometimes lead to the generation of non-words, lemmatization considers the context and converts words to their canonical forms. This process helps in reducing the complexity of the dataset by grouping together various inflections of a word, thereby enhancing the model's ability to learn from the data. Finally, vectorization is the process of converting text into numerical vectors, a format that can be processed by machine learning algorithms. This involves encoding the tokenized and cleaned text into sequences of numbers, where each number corresponds to a word or token in a predefined dictionary. To ensure that these sequences fit into the model uniformly, padding is applied where necessary, standardizing the length of the input sequences across the dataset.

3.2. Model

In the context of emotion recognition from textual data, selecting the appropriate deep learning architecture is crucial to capture the subtleties and complexities of human emotions expressed in language. This study rigorously explores three distinct deep learning models: LSTMs, CNNs, and RNNs, each chosen for its unique capabilities in processing sequential data and its potential to discern emotional undertones in text. LSTMs are a type of RNN specially designed to avoid the long-term dependency problem, making them adept at handling the challenges of sequence prediction tasks where context plays a significant role. The LSTM model in this study is architected to harness this strength, featuring layers that are sequentially arranged to progressively refine the understanding of text data. LSTMs maintain state over long sequences, allowing them to capture not just immediate lexical cues but also the broader context that unfolds over sentences or paragraphs. This is achieved through intricate gating mechanisms comprised of forget, input, and output gates, systematically controlling the flow of information.

The CNN model, traditionally renowned for image processing, has been adapted for text analysis due to its proficiency in detecting local patterns. In text emotion recognition, CNNs scan through word embeddings, capturing pivotal features from fixed-size chunks of the text, which can be indicative of emotional expressions. This model employs layers of convolution followed by pooling to reduce the dimensionality of the data, thus highlighting salient features and patterns in the text that are essential for distinguishing between different emotions. The RNN model excels in processing sequences of data, making it naturally suited for text where the order and flow of words convey meaning. Unlike standard feedforward neural networks, RNNs have loops within them, allowing information to persist. In this study, the RNN model is crafted to sequentially process words, assimilating the emotional context progressively across the text. This continuous integration of context provides a dynamic understanding of the text's emotional trajectory, enabling the model to adjust its predictions based on the flow of text.

All three models are integrated with an Embedding layer at the beginning of the architecture, which transforms words into dense vectors of fixed size, capturing the semantic meaning of words in a lower-dimensional space. This is crucial for the models to interpret text as it converts discrete word tokens into a continuous vector space. Dropout layers are incorporated to prevent overfitting by randomly omitting a subset of features during training, thus ensuring that the model does not rely too heavily on any single element. Bidirectional layers in LSTM and RNN architectures enable the models to process text from both forward and backward directions, enhancing their ability to understand context and reducing the risk of losing information from the end of the sequence. Dense layers follow the recurrent or convolutional layers to perform the final classification, mapping the extracted features to the emotion categories. To facilitate the learning process, the 'adam' optimizer is employed for its efficient gradient descent algorithm, which adjusts the learning rate dynamically, optimizing the network weights and biases. The sparse_categorical_crossentropy loss function is selected due to its compatibility with multi-class classification tasks, measuring the disparity between the predicted emotion distributions and the actual labels, guiding the model toward better accuracy.

3.3. Model Training and Evaluation

In the intricate process of training and evaluating models for emotion recognition from text, a methodical and comprehensive approach is essential to ensure the validity and reliability of the research outcomes. This study adheres to a structured methodology that encompasses model training, robust evaluation through Stratified K-Fold Cross-Validation, meticulous performance assessment using several metrics, and thorough statistical analysis to validate the findings. The models under investigation—LSTMs, CNNs, and RNNs—are trained on a meticulously curated dataset, representative of a broad spectrum of human emotions expressed through text. The dataset is divided into two segments: one portion is utilized for training the models, enabling them to learn and adapt to the patterns and nuances of emotional expression in textual data. The remaining part of the dataset is reserved for testing, serving as a benchmark to evaluate the models' performance and their capability to generalize what they have learned to new, unseen data. To counteract potential biases and ensure the reliability of the evaluation, the study employs Stratified K-Fold Cross-Validation. This technique is particularly effective in addressing imbalances within the class distributions—a common challenge in emotion recognition tasks where some emotions may be more prevalent than others. By preserving the percentage of samples for each class, Stratified K-Fold Cross-Validation enhances the generalizability of the findings, offering a more accurate reflection of the models' performance across diverse emotional states.

The assessment of model performance is comprehensive, leveraging four key metrics: precision, recall, f1-score, and accuracy. Precision measures the proportion of correctly identified positive cases among all cases identified as positive, providing insight into the model's accuracy in predicting each emotion. Recall, or sensitivity, evaluates the model's ability to capture all relevant instances of an emotion, highlighting its sensitivity to the nuances of emotional expression. The f1-score, a harmonic mean of precision and recall, offers a balanced view of the model's performance, considering both the accuracy of predictions and the model's ability to identify all relevant instances. Accuracy, the simplest of the metrics, calculates the proportion of all correct predictions, giving an overall view of the model's effectiveness across all emotional categories. The cross-validation process meticulously partitions the dataset into 'k' equal folds, ensuring that each fold acts as a standalone test set once, while the model is trained on the remaining 'k-1' folds. This cyclic process, repeated 'k' times, allows every data point to be used for both training and testing, thereby maximizing the utility of the dataset. The aggregation of results from each fold culminates in a final performance estimate for each model, offering a robust measure of their ability to recognize and classify emotions accurately.

4. Results and Discussions

The exploration of deep learning models for emotion recognition in text has unveiled insightful findings that shed light on the intricacies of natural language processing and the nuanced detection of emotional undertones. Through a detailed examination of LSTMs, CNNs, and RNNs, this study offers a comprehensive analysis of each model's efficacy in deciphering the complex web of human emotions expressed through language. As presented in table 1, the standout performer in this study, LSTMs, exhibited remarkable proficiency, with a mean accuracy rate of 93.48%. This high level of accuracy is complemented by a precision of 93.70%, indicating the LSTM model's adeptness at minimizing false positives in emotion classification—a critical factor in applications where misinterpretation can have significant repercussions, such as mental health assessment or customer sentiment analysis. The recall rate and F1 score further corroborate the LSTM's balanced capability, ensuring that it not only identifies emotions with high accuracy but also does so consistently across the spectrum of emotional expressions. This can be attributed to the LSTM's architectural design, which integrates memory cells that effectively capture and utilize long-term contextual information, making it exceptionally suited for tasks where the sequence and flow of text significantly influence meaning.

Table 1. Deep Learning Performance

Models	**Accuracy**	**Precision**	**Recall**	**F1 Score**
LSTMs	0.93483346	0.9369998	0.934833460	0.9344546
CNN	0.91782087	0.9180185	0.917820875	0.9171164
RNN	0.44934703	0.2702387	0.449347036	0.3155029

In comparison, CNNs demonstrated considerable prowess with a mean accuracy of 91.78%, albeit slightly behind LSTMs. This performance is indicative of CNNs' strength in extracting pivotal features from fixed-size segments of text, which can be particularly beneficial for identifying specific emotional cues or patterns within data. The precision and recall rates suggest a high degree of model reliability, with the F1 score reinforcing CNNs' role as a potent model for emotion recognition. However, the slightly diminished performance compared to LSTMs may reflect the inherent limitation of CNNs in fully capturing the sequential and temporal dynamics of textual data, as they primarily focus on local dependencies. Conversely, the performance of RNNs in this context was markedly lower, with accuracy at 44.93%, precision at 27.02%, and an F1 score of 31.55%. These figures highlight the fundamental challenges that basic RNN architectures face in processing complex linguistic structures and maintaining contextual information over long text sequences. The inherent limitations of RNNs, including difficulties with long-term dependencies and susceptibility to the vanishing gradient problem, are likely contributors to their subpar performance. This underscores the necessity for more sophisticated or specialized recurrent models that can navigate the complexity of emotion recognition tasks more effectively.

The implications of these findings extend far beyond the comparative performance of the models. The superior results of LSTMs underscore the importance of context and memory in understanding emotional nuances in text, suggesting that future research should further explore and refine these aspects of model architecture. Additionally, the respectable performance of CNNs highlights the potential for hybrid models that combine local feature extraction with sequential processing, offering a promising direction for enhancing emotion recognition capabilities. The underwhelming performance of traditional RNNs raises critical questions about the evolution of model architectures and the continuous search for more efficient and accurate methods of processing sequential data. It prompts a reassessment of the roles different models play in the broader landscape of NLP and emotion recognition, advocating for a nuanced approach that tailors the model selection to the specific characteristics of the task at hand.

5. Conclusion

This research ventured into the realm of emotion recognition from textual data, focusing on the comparative analysis of three prominent deep learning models: LSTMs, CNNs, and RNNs. Through the meticulous application of 5-fold cross-validation, the study provided a clear and detailed performance assessment, revealing significant differences in the models' ability to accurately classify emotions. LSTMs emerged as the superior model, demonstrating exceptional accuracy, precision, recall, and F1 score, indicative of their robust capability to capture and interpret the nuanced and contextually rich nature of emotional text. This finding underscores the critical importance of considering temporal dependencies and contextual nuances in text-based emotion recognition, areas where LSTMs inherently excel. Furthermore, CNNs, while slightly lagging LSTMs in performance metrics, still showcased a strong ability to recognize emotional patterns in text. Their efficiency in feature extraction from textual data positions them as a valuable tool for emotion recognition tasks, particularly when dealing with large datasets or requiring rapid processing. RNNs, in their basic form, displayed limitations, struggling to match the performance of their more advanced counterparts.

This outcome highlights the challenges faced by simpler recurrent models in handling the complexities of language and emotion, suggesting a need for advancements or modifications in architecture to improve their efficacy in such tasks. The study's findings contribute significantly to the field of natural language processing and emotion recognition, providing clear evidence of the varying strengths and weaknesses of different deep learning models in this context. It also opens new avenues for future research, particularly in exploring hybrid models or advanced architectures like attention mechanisms, which could potentially combine the strengths of LSTMs and CNNs for even more accurate emotion detection. Moreover, this research has practical implications for a wide array of applications, from enhancing user experience in digital communications to supporting mental health assessments through sentiment analysis. The insights gained from this study can guide the development of more sensitive, accurate, and nuanced emotion recognition systems, ultimately leading to advancements in human-computer interaction and artificial intelligence.

References

[1] P. Nandwani and R. Verma, “A review on sentiment analysis and emotion detection from text,” Soc. Netw. Anal. Min., vol. 11, no. 1, p. 81, 2021.

[2] A. Ramsay and T. Ahmad, Machine Learning for Emotion Analysis in Python: Build AI-powered tools for analyzing emotion using natural language processing

and machine learning. Packt Publishing Ltd, 2023.

[3] S. Kusal, S. Patil, J. Choudrie, K. Kotecha, D. Vora, and I. Pappas, "A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection," Artif. Intell. Rev., vol. 56, no. 12, pp. 15129–15215, 2023.

[4] W. Chung and D. Zeng, "Dissecting emotion and user influence in social media communities: An interaction modeling approach," Inf. & Manag., vol. 57, no. 1, p. 103108, 2020.

[5] K. Chakraborty, S. Bhattacharyya, and R. Bag, "A survey of sentiment analysis from social media data," IEEE Trans. Comput. Soc. Syst., vol. 7, no. 2, pp. 450–464, 2020.

[6] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, "Text-based emotion detection: Advances, challenges, and opportunities," Eng. Reports, vol. 2, no. 7, p. e12189, 2020.

[7] D. Kalla and N. Smith, "Study and analysis of chat GPT and its impact on different fields of study," Int. J. Innov. Sci. Res. Technol., vol. 8, no. 3, 2023.

[8] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," Artif. Intell. Rev., vol. 53, no. 6, pp. 4335–4385, 2020.

[9] C. Dalvi, M. Rathod, S. Patil, S. Gite, and K. Kotecha, "A survey of ai-based facial emotion recognition: Features, ml & dl techniques, age-wise datasets and future directions," Ieee Access, vol. 9, pp. 165806–165840, 2021.

[10] H. I. Abdalla and A. A. Amer, "On the integration of similarity measures with machine learning models to enhance text classification performance," Inf. Sci. (Ny)., vol. 614, pp. 263–288, 2022.

[11] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Deep Learning applications for COVID-19," J. big Data, vol. 8, no. 1, pp. 1–54, 2021.

[12] G. Chandrasekaran, S. Dhanasekaran, C. Moorthy, and A. Arul Oli, "Multimodal sentiment analysis leveraging the strength of deep neural networks enhanced by the XGBoost classifier," Comput. Methods Biomech. Biomed. Engin., pp. 1–23, 2024.

[13] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," Artif. Intell. Rev., vol. 53, no. 8, pp. 5929–5955, 2020.

[14] M. Rai and J. K. Pandey, Using Machine Learning to Detect Emotions and Predict Human Psychology. IGI Global, 2024.

[15] A. Onan and K. F. Balbal, "Improving Turkish text sentiment classification through task-specific and universal transformations: an ensemble data augmentation approach," IEEE Access, 2024.

[16] F. Carvalho, F. P. Junior, E. Ogasawara, L. Ferrari, and G. Guedes, "Evaluation of the Brazilian Portuguese version of linguistic inquiry and word count 2015 (BP-LIWC2015)," Lang. Resour. Eval., vol. 58, no. 1, pp. 203–222, 2024.

[17] F. Mehraliyev, I. C. C. Chan, and A. P. Kirilenko, "Sentiment analysis in hospitality and tourism: a thematic and methodological review," Int. J. Contemp. Hosp. Manag., vol. 34, no. 1, pp. 46–77, 2022.

[18] S. Alshattnawi, A. Shatnawi, A. M. R. AlSobeh, and A. A. Magableh, "Beyond Word-Based Model Embeddings: Contextualized Representations for Enhanced Social Media Spam Detection," Appl. Sci., vol. 14, no. 6, p. 2254, 2024.

[19] L. Williams, E. Anthi, and P. Burnap, "Comparing hierarchical approaches to enhance supervised emotive text classification," Big Data Cogn. Comput., vol. 8, no. 4, p. 38, 2024.

[20] S. Lee, D. K. Han, and H. Ko, "Fusion-ConvBERT: parallel convolution and BERT fusion for speech emotion recognition," Sensors, vol. 20, no. 22, p. 6688, 2020.

[21] J. Kocoń et al., "ChatGPT: Jack of all trades, master of none," Inf. Fusion, vol. 99, p. 101861, 2023.
[22] Y. Zhu et al., "Large language models for information retrieval: A survey," arXiv Prepr. arXiv2308.07107, 2023.
[23] P. H. Andersen et al., "Towards machine recognition of facial expressions of pain in horses," Animals, vol. 11, no. 6, p. 1643, 2021.
[24] S. Bursic and others, "On Wiring Emotion to Words: A Bayesian Model," 2022.
[25] C. Thaokar, J. K. Rout, M. Rout, and N. K. Ray, "N-Gram Based Sarcasm Detection for News and Social Media Text Using Hybrid Deep Learning Models," SN Comput. Sci., vol. 5, no. 1, p. 163, 2024.
[26] S. L. Heggernes, "A critical review of the role of texts in fostering Intercultural Communicative competence in the English Language classroom," Educ. Res. Rev., vol. 33, p. 100390, 2021.
[27] Q. C. Song, C. Tang, and S. Wee, "Making sense of model generalizability: A tutorial on cross-validation in R and Shiny," Adv. Methods Pract. Psychol. Sci., vol. 4, no. 1, p. 2515245920947067, 2021.
[28] Kotholo, "Emotions Logistic Regression 93%." 2020.