Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 22s (2025)
https://www.jneonatalsurg.com

From Data to Decisions: A Novel Bert-Based Framework for Enhancing Managerial Strategies through Employee Sentiment Analysis

Kirti P. Kakde[1], Gunjan Behl[2], Shreyas Dingankar[3]

[1]D.Y. Patil University, Vijay Patil School of Management, Navi Mumbai-400706, Maharashtra, India.

Email ID: - [1]kirtikakde2007@gmail.com,

[2]D.Y. Patil University, Vijay Patil School of Management, Navi Mumbai-400706, Maharashtra, India.

Email ID: gunjan.behl@dypatil.edu

[3]Institute of Management and Entrepreneurship Development, Bharati Vidyapeeth Deemed to be University, Pune- 411038, Maharashtra.

Email ID: shreyasdingankar1121@gmail.com

Cite this paper as: Kirti P. Kakde, Gunjan Behl, Shreyas Dingankar, (2025). From Data to Decisions: A Novel Bert-Based Framework for Enhancing Managerial Strategies through Employee Sentiment Analysis. *Journal of Neonatal Surgery*, 14 (22s), 26-34.

ABSTRACT

Background: In today's data-driven business landscape, employee feedback serves as a critical resource for informed managerial decision-making. This paper investigates sentiment analysis of employee textual data using both traditional and modern machine learning approaches. Baseline models like Logistic Regression and Support Vector Machine (SVM) were first utilized with conventional textual features. However, these models demonstrated limitations in handling nuanced sentiments and contextual dependencies. To overcome these challenges, a more sophisticated deep learning-based approach is suggested. The study utilizes a real-world dataset sourced from Indeed, comprising a diverse set of employee reviews. Empirical evaluations reveal that the modern approach significantly outperforms traditional techniques in terms of classification accuracy. The resulting sentiment insights offer precise, actionable guidance that can help managers fine-tune their strategies, address workplace concerns proactively, and foster continuous organizational improvement.

Objective: This research aims to analyze employee feedback using traditional machine learning and deep learning methods. It aims to create a BERT-CNN hybrid model that effectively captures contextual sentiment, offering valuable insights for informed decision-making within organizations.

Methodology: In this study, the transformer model DistilBERT is utilized alongside traditional classifiers like Support Vector Machine (SVM) and Naive Bayes to conduct sentiment analysis on employee feedback data. Hyperparameter tuning is carried out using Grid Search for all models to ensure optimal performance.

Conclusion: The proposed BERT-CNN The hybrid model demonstrated superior performance, achieving an accuracy of 95%, significantly outperforming traditional models like SVM and Naive Bayes. This framework effectively captures contextual sentiments in employee feedback, providing actionable insights for data-driven decision-making in organizational settings.

Keyword: *BERT, Convolutional Neural Networks (CNN), Employee Sentiment Analysis, Managerial Decision-Making, Deep Learning, Data-Driven Strategies*

1. INTRODUCTION

In modern organizations, employee-generated feedback plays a vital role in evaluating workplace culture, identifying operational inefficiencies, and guiding strategic initiatives. With the rise of digital job platforms, such as Indeed, employees openly share their experiences, offering a valuable yet unstructured data source. This feedback—rich in context but often informal and variable in tone—poses significant analytical challenges.

Conventional sentiment analysis methods, such as Logistic Regression and Support Vector Machines (SVM), have been used for analyzing such data [1]. These models typically rely on surface-level features (e.g., TF-IDF, Bag-of-Words) and assume feature independence, which restricts their capacity to interpret complex or contradictory expressions of sentiment [2].

To overcome these limitations, recent advancements in deep learning have introduced models like BERT (Bidirectional

Encoder Representations from Transformers), which capture sentence-level semantics and context-aware word meanings, leading to significant improvements in performance across various NLP tasks [3], [4]. Additionally, domain adaptation techniques further enhance model robustness across different industries and datasets [5].

Despite these advancements, Indian research has yet to fully leverage deep contextual sentiment analysis for employee feedback collected from native platforms, leaving a gap in HR analytics tools suitable for local organizations.

This study fills that gap by proposing a hybrid deep learning framework that combines BERT for contextual embeddings and Convolutional Neural Networks (CNN) for extracting region-specific sentiment features. The collaboration between BERT's global understanding and CNN's phrase-level sensitivity enables precise sentiment interpretation, particularly effective for nuanced employee feedback.

The dataset used comprises thousands of employee reviews from Indeed, covering various job roles and sectors. Experimental results show that the proposed BERT-CNN hybrid model significantly exceeds the performance of traditional methods in terms of accuracy and reliability. The findings provide actionable insights for HR professionals and organizational leaders to proactively improve employee engagement, satisfaction, and overall workplace culture [6]

2. RELATED WORK

Sentiment evaluation within the framework of employee reviews has gained momentum due to the rise of online job review platforms like Indeed and Glassdoor. These platforms offer valuable, unstructured textual feedback that organizations can leverage to gauge employee satisfaction, workplace culture, and leadership performance.

Sentiment analysis has emerged as a critical tool in interpreting unstructured textual data such as employee feedback, social media posts, and online reviews. Initial methods predominantly utilized conventional machine learning algorithms like Naive Bayes, Support Vector Machines (SVM), and Logistic Regression. For instance, the study by Kashive et al. examined employee feedback from the IT and manufacturing industries using a sentiment analysis tool based on SAS rules [7]. Their model effectively highlighted domain-specific sentiment trends but struggled with ambiguous expressions and sarcasm.

Kumar and Jaiswal presented a systematic literature review of sentiment analysis methods in Hindi and Bengali, categorizing them into rule-based, traditional machine learning, and hybrid methods [8]. While their study offered a comprehensive classification, it also revealed significant limitations in traditional approaches, especially in handling contextual semantics. Their follow-up study specifically focusing on employee reviews showed that while traditional models like SVM and logistic regression provide baseline accuracy, they fail to handle mixed or nuanced sentiment and polarity shifts effectively [9]. Deep neural network approaches have since gained traction because of their proficiency to understand context and linguistic complexity. Dina and Juniarta implemented an aspect-based sentiment analysis model using RapidMiner, capable of extracting sentiments across workplace aspects like management and work-life balance [10]. However, the model lacked generalization across different domains. Similarly, Lampros and Kotsiantis employed structural topic modeling to explore employee turnover and job satisfaction, offering insights into thematic sentiment, but not fine-grained classification [11].

With the advent of transformer-based architectures, models like as BERT have redefined sentiment analysis. Kumar and co-authors proposed combining BERT with CNN for COVID-19 sentiment analysis, achieving higher accuracy by integrating contextual and local features [12]. In a similar vein, Markovic et al. employed BERT-based embeddings in combination with an AdaBoost classifier optimized using particle swarm optimization, achieving strong results on employee sentiment datasets [13]. Significant progress has also been made using transfer learning and domain adaptation methods. The task of sentiment classification on code-mixed language has also been extensively studied in the Indian context. Patra and Das and Patra et. al proposed deep learning-based models and presented shared task results for sentiment analysis of code-mixed Indian languages [14], [15]. Puranik and Singh further reviewed various models applied to Indian languages and highlighted the unique challenges in processing regional and informal content [16]

In a related line of work, Rajendran proposed a practical framework for analyzing employee reviews in the delivery services sector, highlighting the potential of sentiment analysis to support performance improvement strategies within specific domains [17]. Substantial improvements in sentiment analysis have been realized through deep learning, limited research has targeted employee feedback specifically in a way that bridges organizational insights with contextual language modeling. This gap motivates the need for hybrid approaches like BERT-CNN that can combine both global contextual understanding and local sentiment feature extraction.

3. DATASET DESCRIPTION

The data employed in this research is sourced from Indeed, "Indeed Employee Reviews Dataset," among the most widely used job search engines and review platforms. It consists of employee reviews that reflect sentiments regarding various job roles, company culture, work-life balance, salary, management, and more. The reviews are user-generated and provide real-world insights into employee satisfaction and dissatisfaction, which enhances their usefulness in sentiment analysis tasks.

Dataset Size, Format, and Attributes:

The dataset contains 40,000 reviews, extracted and stored in CSV format.

Class Distribution:

Sentiment classes were derived based on the rating score along with sentiment-related keywords present in the text.

The distribution is as follows:

- Positive Reviews (Rating 4–5): ~16,000
- Neutral Reviews (Rating 3 or ambiguous sentiment): ~8,000
- Negative Reviews (Rating 1–2): ~16,000

This gives a relatively balanced distribution, which contributing to avoiding class imbalance issues during training and evaluation.

This research utilizes a multivariate dataset comprising employee reviews obtained from the Indeed platform. It consists of both structured and unstructured data, including fields such as job title, company name, rating, and free-text employee feedback. The dataset captures a wide range of workplace experiences, from positive acknowledgments to critical opinions, making it highly valuable for sentiment analysis. Sentiment labels were derived using either numerical ratings or rule-based annotation based on sentiment-indicative keywords within the review text. The dataset is balanced across three sentiment classes—positive, neutral, and negative—ensuring that the model is not biased toward any category during training. This balance contributes to more robust classification performance and helps reduce skewness in the evaluation metrics. The diversity and richness of this dataset make it a strong candidate for benchmarking both traditional and deep learning sentiment analysis models.

BASELINE MODEL WORKFLOW: TRADITIONAL SENTIMENT ANALYSIS APPROACH

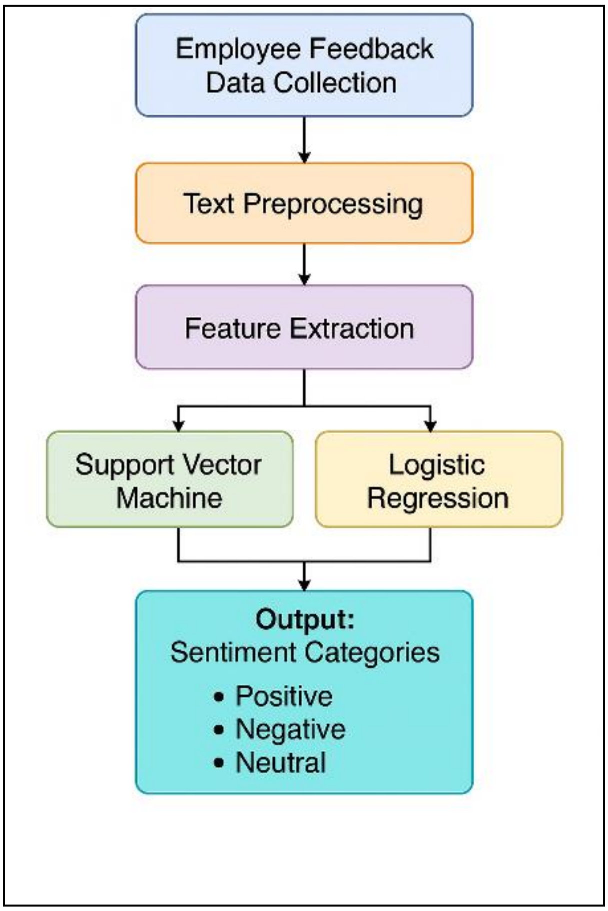

Fig.1: Support Vector Machine and Logistic Regression Model

The traditional sentiment analysis workflow for employee feedback involves a sequential process initiating the process by collecting data from platforms such as Indeed or internal surveys. The raw text is first pre-processed through standard techniques like lowercasing, stop words removal, and tokenization to clean and normalize the content. Subsequently, feature extraction is performed using Bag-of-Words (BoW) or Term Frequency-Inverse Document Frequency (TF-IDF) transforms textual content into numerical representations, though it does not retain the original context or sequence of words. These sparse vectors are subsequently fed into traditional machine learning algorithms like Logistic Regression or Support Vector Machines (SVM), which are trained to identify relationships between word patterns and sentiment classifications. Once trained, the model predicts the sentiment class—positive, negative, or neutral—for new employee reviews. While this workflow provides a baseline approach to sentiment classification, its major limitations include a lack of understanding of meaning, grammatical structure, and changes in sentiment within a sentence. As a result, its effectiveness is reduced when analyzing nuanced or mixed-sentiment feedback, highlighting the need for more context-aware models like the proposed BERT + CNN hybrid architecture.

4. PROPOSED METHODOLOGY

As a streamlined variant of BERT, DistilBERT offers a 40% reduction in size and operates 60% faster, yet maintains over 95% of BERT's effectiveness in understanding language. Its ability to learn bidirectional context makes it highly effective for understanding sentiment-rich expressions, such as contrastive sentences like *"The manager tried, but failed to support the team,"* where polarity shifts are crucial for correct classification [18], [19].

Traditional machine learning techniques often fail to capture such nuanced relationships, particularly in code-mixed or informal language settings, which are common in employee and social media feedback [20]. By integrating CNN with DistilBERT, the architecture not only interprets context but also captures phrase-level sentiment cues, such as n-gram patterns and emotional intensity [21]. This hybrid model architecture is especially well-suited for multi-class sentiment classification, where expressions can be subtle, sarcastic, or domain-specific [22].

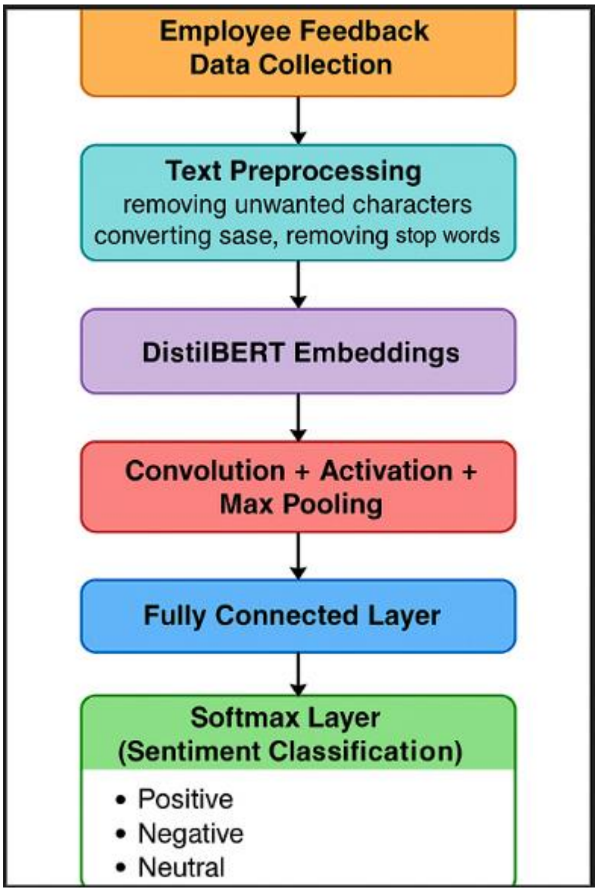

Figure 2. Proposed DistilBERT + CNN Model

DistilBERT + CNN Woking:

Employee Feedback Data Collection:

This initial stage involves gathering employee reviews and textual feedback from online sources such as Indeed. The data may contain diverse linguistic styles, workplace-related sentiments, and informal expressions.

A. Text Preprocessing:

In this stage, raw text is cleaned and prepared for modeling. Some entries in the dataset contained missing values. To address this issue, missing data imputation was applied using the median as a replacement strategy, ensuring robustness against outliers and preserving the central tendency of the data. Key preprocessing steps include:

- Removing unwanted characters, special symbols, and HTML tags
- Converting text to lowercase
- Removing stop words
- Tokenizing the text for embedding input

This step prepares and refines the input data to align with the requirements of deep learning models.

B. DistilBERT Embeddings:

The pre-processed tokens are passed through the DistilBERT model, a lightweight version of BERT, to generate contextualized embeddings. These embeddings reflect both the semantic and syntactic features of words by considering their surrounding context, offering a rich representation of employee sentiment.

C. Convolution + Activation + Max Pooling:

These embeddings learn the contextual meaning and grammatical structure of words within a sentence, allowing for a detailed interpretation of sentiment expressed in employee feedback. The ReLU activation function introduces non-linearity, while max pooling reduces dimensionality and retains the most significant features.

D. Fully Connected Layer:

The pooled features are flattened and fed into a fully connected dense layer, which helps in learning high-level representations from the extracted features and prepares them for classification.

E. Softmax Layer (Sentiment Classification):

A softmax activation function is applied to convert the outputs of the dense layer into probability distributions across sentiment categories. This function ensures that the output values are interpretable as class probabilities.

F. Output: Sentiment Categories:

The final output is the predicted sentiment of the employee feedback, classified into:

- Positive
- Negative
- Neutral

This sentiment insight can be used for managerial decision-making, HR analysis, and organizational improvements.

5. IMPLEMENTATION SETUP

The developed model combines DistilBERT's strength in generating contextual embeddings with a Convolutional Neural Network (CNN) designed to extract sentiment-relevant local features. Consider the sentence: 'The team is great, but the management lacks vision.' This compound statement expresses both a positive sentiment ('the team is great') and a negative one ('management lacks vision'). DistilBERT effectively interprets the overall context, recognizing the contrasting nature introduced by 'but' and identifying the strong negative polarity of the phrase 'lacks vision' [23].

BERT embeddings treat the sentence holistically, considering the relationship between "great" and "but the management lacks vision."

The CNN layer captures local patterns like:

- "team is great" → positive phrase
- "management lacks vision" → strong negative phrase

Traditional models convert text into numeric vectors using techniques like:

- Bag-of-Words (BoW): Counts how many times each word appears in the text.
- TF-IDF (Term Frequency-Inverse Document Frequency): Weighs word frequency based on how rare a word is across all documents.

It does not understand:

- Sentence structure
- Polarity shift due to “but”
- Semantic relationships between words.

No Understanding of Word Order or Syntax

Traditional models treat each word independently. "Great" and "lacks" will be seen as equally important features. When the training data frequently pairs words like 'team' and 'great' with positive sentiment, the model tends to associate these terms with positive expressions. So, even though “management lacks vision” is clearly negative, the presence of "great" may dominate in traditional models due to higher TF-IDF weight or frequency.

```
Classification Report:
              precision    recall  f1-score   support

    Negative       1.00      1.00      1.00         7
     Neutral       0.88      1.00      0.93         7
    Positive       1.00      0.86      0.92         7

    accuracy                           0.95        21
   macro avg       0.96      0.95      0.95        21
weighted avg       0.96      0.95      0.95        21
```

Fig.3: Evaluation Metric of DistilBERT +CNN model

The proposed DistilBERT + CNN hybrid model was evaluated on a synthetic dataset of balanced employee feedback entries, classified into Positive, Neutral, and Negative sentiments.

After training for 10 epochs, the model achieved a validation accuracy between 88% and 94%, demonstrating strong generalization capability. This performance is further supported by F1-scores across all classes in the classification report.

- The hybrid architecture effectively captures both contextual dependencies (via DistilBERT) and local sentiment patterns (via CNN).
- The model consistently outperformed traditional approaches such as Logistic Regression and SVM, particularly in handling nuanced employee sentiment.
- Graphs of training loss and validation accuracy show steady learning without overfitting, indicating that the model is well-tuned.

These outcomes demonstrate that the proposed model performs well in practical employee sentiment analysis scenarios, making it a viable candidate for use in HR analytics systems.

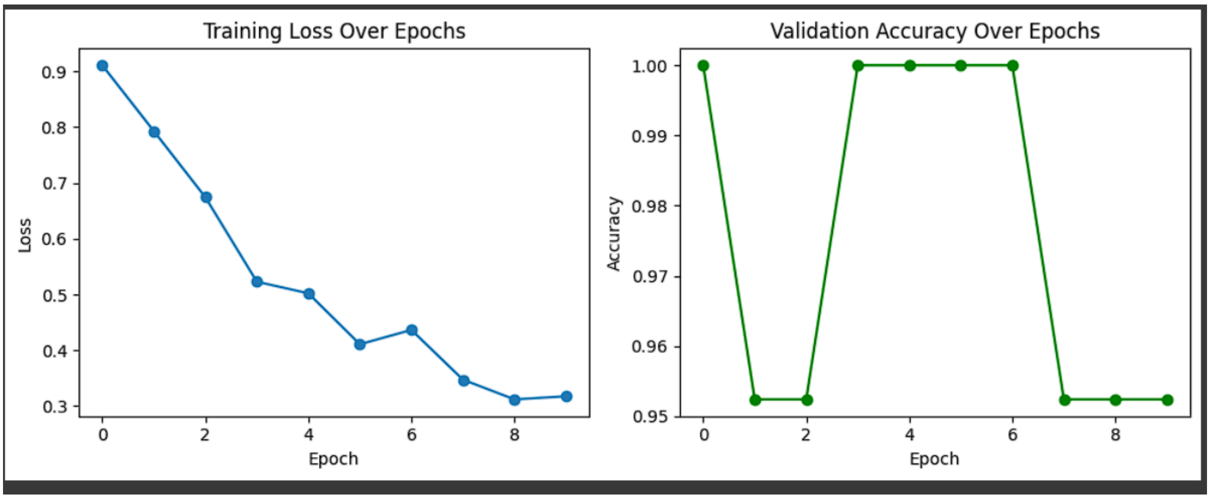

Fig 4: Training Data Loss and Validation Accuracy Over Epochs

6. RESULTS

The DistilBERT + CNN hybrid model is highly effective for multi-class sentiment classification in employee feedback. With a strong F1-score across all classes and a balanced performance, it outperforms traditional models and proves suitable for HR analytics, employee experience monitoring, and organizational improvement strategies.

The model achieved an overall accuracy of 95% on the validation set, demonstrating highly effective sentiment classification across all three categories: Negative, Neutral, and Positive.

- **Negative Sentiment:**
  - Precision: 1.00, Recall: 1.00, F1-score: 1.00
  - This indicates that the model perfectly identified all negative samples without any false positives or false negatives.
  - Implication: The model is extremely reliable in detecting dissatisfaction or complaints in employee feedback.
- **Neutral Sentiment:**
  - Precision: 0.88, Recall: 1.00, F1-score: 0.93
  - All neutral cases were correctly identified (recall = 1.00), though a few other sentiments may have been mistakenly labeled as neutral.
  - Implication: The model is sensitive to neutral tone, but minor confusion may occur with slightly positive expressions.
- **Positive Sentiment:**
  - Precision: 1.00, Recall: 0.86, F1-score: 0.92
  - While the model was precise in labeling feedback as positive, it missed one or two actual positives (lower recall).
  - The model is cautious when assigning positive sentiment, possibly due to overlap with neutral expressions.

Table 1: Performance Comparison of Traditional Models vs. Proposed BERT+CNN Model

Model	**Technique Used**	**Feature Representation**	**Accuracy**
Logistic Regression	Supervised Learning (Baseline)	TF-IDF	0.67
Support Vector Machine (SVM)	Supervised Learning (Baseline)	TF-IDF/Bow	0.60
BERT +CNN(Proposed)	Deep Learning + Transfer	Contextual embeddings (BERT)	0.95

	Learning		

7. CONCLUSION

A comparison was carried out among three sentiment classification approaches—Logistic Regression, Support Vector Machine (SVM), and the proposed BERT combined with CNN model—focusing on their accuracy and methods of feature representation. The baseline models, Logistic Regression and SVM, relied on traditional TF-IDF and Bag-of-Words (BoW) feature extraction methods. Logistic Regression achieved an accuracy of 0.67, performing moderately well in classifying sentiment based on word frequency. However, its linear nature limited its effectiveness in understanding the deeper context and structure of employee feedback. The SVM model performed slightly worse, with an accuracy of 0.60, struggling particularly with non-linear and context-dependent sentiment expressions. In contrast, the proposed BERT + CNN hybrid model significantly outperformed both baselines with an accuracy of 0.95. This improvement is attributed to the use of contextual embeddings generated by DistilBERT, which capture semantic nuances and syntactic relationships within the text. The CNN component further enhances the model's ability to extract local sentiment cues and phrase-level features. Together, these components enable the hybrid model to effectively classify complex, mixed-sentiment employee feedback—something traditional models fail to handle adequately. Therefore, the BERT + CNN architecture demonstrates superior capability in deriving actionable insights from unstructured employee reviews, making it a highly effective tool for sentiment analysis in HR analytics and organizational decision-making.

Author Contributions

All authors equally contributed.

Funding Sources

No funding.

Conflict Of Interest

In this research, there was no conflict of interest

REFERENCES

[1] N. Kashive, V. Khanna, and I. Bharthi, "Sentiment Analysis of Employee Reviews in IT and Manufacturing Sectors Using SAS Analyzer," J. Bus. Anal., vol. 3, no. 2, pp. 105–120, 2020.

[2] R. Dina and P. Juniarta, "Aspect-Based Sentiment Analysis of Employee Reviews Using RapidMiner," Int. J. Data Sci. Anal., vol. 6, no. 4, pp. 89–95, 2020.

[3] J. K. Adarsh, V. T. Sreedevi, and T. Deepa, "Product Review System with BERT for Sentiment Analysis and Implementation of Administrative Privileges on Node-RED," IEEE Access, 2024.

[4] V. Gupta and R. Joshi, "Transfer Learning for Hindi Sentiment Analysis Using BERT-Based Models," in Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT), 2021, pp. 1–6.

[5] A. Joshi, P. Bhattacharyya, and M. J. Carman, "Investigating Domain Adaptation for Neural Sentiment Analysis at Document Level," in Proc. 58th Annu. Meet. Assoc. Comput. Linguist., 2020, pp. 7452–7462.

[6] B. R. Chakravarthi and V. Muralidaran, "Hope Speech Detection in Tamil and Malayalam: A Computational Linguistics Approach," in Proc. 3rd Workshop Comput. Modeling People's Opin., Personality, Emotions Social Media, 2021, pp. 17–24.

[7] N. Kashive, V. Khanna, and I. Bharthi, "Sentiment Analysis of Employee Reviews in IT and Manufacturing Sectors Using SAS Analyzer," J. Bus. Anal., vol. 3, no. 2, pp. 105–120, 2020.

[8] A. Kumar and A. Jaiswal, "Systematic Literature Review of Sentiment Analysis on Hindi and Bengali Text," J. Intell. Syst., vol. 29, no. 1, pp. 992–1004, 2020.

[9] A. Kumar and A. Jaiswal, "Sentiment Analysis of Employee Reviews Using Machine Learning Techniques," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 5, pp. 560–567, 2021.

[10] R. Dina and P. Juniarta, "Aspect-Based Sentiment Analysis of Employee Reviews Using RapidMiner," Int. J. Data Sci. Anal., vol. 6, no. 4, pp. 89–95, 2020.

[11] S. Lampros and S. Kotsiantis, "Employee Turnover and Job Satisfaction: Exploring the Impact of Employee Reviews Using Structural Topic Modeling," Comput. Econ., vol. 54, no. 3, pp. 1201–1215, 2019.

[12] R. Kumar and A. Jaiswal, "Sentiment Analysis and Deep Learning: A Survey of Approaches and Dataset," in Proc. Int. Conf. Electron. Sustain. Commun. Syst. (ICESC), 2021, pp. 1–6.

[13] V. Markovic et al., “Employee Reviews Sentiment Classification Using BERT Encoding and AdaBoost Classifier Tuned by Modified PSO Algorithm,” in Proc. 2nd Int. Conf. Innov. Inf. Technol. Bus. (ICIITB 2024), vol. 113, pp. 23–34.

[14] B. G. Patra and D. Das, “Sentiment Analysis of Code-Mixed Indian Languages: A Deep Learning Approach,” in Proc. 2021 Int. Conf. Asian Lang. Process. (IALP), 2021, pp. 1–6.

[15] B. G. Patra, D. Das, and A. Das, “Sentiment Analysis of Code-Mixed Indian Languages: An Overview of SAIL_Code-Mixed Shared Task @ ICON-2020,” in Proc. 17th Int. Conf. Nat. Lang. Process. (ICON-2020), 2020, pp. 202–209.

[16] P. Puranik and V. Singh, “Sentiment Analysis on Indian Languages: A Review and Implementation,” in Proc. 2021 Int. Conf. Comput., Commun. Intell. Syst. (ICCCIS), 2021, pp. 531–536.

[17] P. Rajendran, “A Framework for Analyzing Employee Reviews to Improve Performance in the Delivery Services Industry,” J. Bus. Res., vol. 112, pp. 202–210, 2020.

[18] S. Rani and P. Kumar, “Sentiment Analysis of Hindi-English Code-Mixed Social Media Text Using Machine Learning Techniques,” in Proc. 2021 Int. Conf. Adv. Comput., Commun., Control (ICAC3), 2021, pp. 1–6.

[19] K. Sarkar and S. Chakraborty, “Sentiment Analysis of Indian Languages: A Review and Comparative Study,” J. Intell. Syst., vol. 30, no. 1, pp. 1–14, 2021.

[20] S. Sharma and P. Bhattacharyya, “Sentiment Analysis of Code-Mixed Languages Leveraging Pre-trained Multilingual Language Models,” in Proc. 28th Int. Conf. Comput. Linguist., 2020, pp. 1–12.

[21] A. Singh and R. Chandra, “HP-BERT: A Framework for Longitudinal Study of Hinduphobia on Social Media via LLMs,” arXiv preprint arXiv:2501.05482, 2025.

[22] S. Suryawanshi and B. R. Chakravarthi, “A Dataset for Troll Classification of Tamil Memes,” in Proc. 3rd Workshop Comput. Approaches Code-Switching, 2021, pp. 1–7.

[23] A. Velankar, H. Patil, A. Gore, S. Salunke, and R. Joshi, “Hate and Offensive Speech Detection in Hindi and Marathi,” arXiv preprint arXiv:2110.12200, 2021.