International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V73I6P113

Volume 73 Issue 6, 104-111, June 2025

Original Article

Hybrid Machine Learning and NLP Approaches for Sentiment Productivity Analysis of Employees

Deepthi M. Pisharody[1], Anjali E.S[2], Maidhili Mohan K[3]

[1,2,3]*Department of Computer Science, Prajyoti Niketan College, Kerala, India.*

[1]*Corresponding Author : deepthi.saji@gmail.com*

Received: 03 May 2025 Revised: 04 June 2025 Accepted: 21 June 2025 Published: 30 June 2025

Abstract *- The active involvement and well-being of employees have an immense effect on organizational productivity. However, traditional review procedures for employee engagement often rely on manual analysis of qualitative feedback, which makes them ineffective and biased. In order to improve the precision and effectiveness of employee assessments, this study suggests an automated framework that combines Machine Learning (ML) and Natural Language Processing (NLP). The system associates the structured performance and the attendance data with textual feedback submitted by employees to show a comprehensive analysis of employee happiness. Prior to TF-IDF vectorisation, text data is pre-processed using common NLP techniques like tokenisation, stemming, and stop-word removal. To sort out the feedback as positive, neutral, or negative, we have used machine learning models such as Random Forest, Naïve Bayes, Support Vector Machine (SVM), and Logistic Regression. Logistic Regression outperformed other machine learning models with an accuracy of 99.01%. Clustering of employees is performed according to the key performance indicators using the K-Means clustering algorithm, which opens up the trends in organizational productivity and employee engagement. The suggested system aligns with modern methodologies that support a comprehensive perspective of employee feedback by integrating sentiment classification, predictive modeling, and clustering into a single pipeline. Linking clustered behaviours to productivity and attendance trends goes beyond simple sentiment polarity and enables organizations to pinpoint not only disgruntled individuals but also systemic problems across departments.*

Keywords *- Artificial Intelligence, K-means clustering, Natural Language Processing, Machine Learning, Employee happiness.*

1. Introduction

In today's competitive business world, productivity, innovation, teamwork, and employee satisfaction demand extra care beyond financial benefits. An organization needs employees who are more loyal, dedicated, and in line with the goals of the organization. But keeping people happy at work and in the workplace still remains a tedious task, and traditional feedback methods used by employers, like annual surveys and periodic evaluations, do not always provide useful or on-time information. The evolution of Machine Learning (ML) and Artificial Intelligence (AI) facilitates a new path to the solution. Combined data from a variety of sources, such as employee feedback, performance metrics, attendance records, and communication patterns, is used by organizations or employers to get complete, real-time pictures of the feelings and reactions of employees.

Using Natural Language Processing (NLP) and sentiment analysis, the emotional subtleties in employee feedback that show stress, unhappiness, or motivation can be identified. The clustering algorithms allow people to be divided into clusters based on their level of engagement and productivity, which helps in the effective customization of HR strategies for each person. Predictive analytics can also detect early signs of burnout or disengagement, which assists companies in stepping in early and building a stronger, more motivated, and more resilient workforce.

2. Related Work

Sentiment analysis, or opinion mining, has become a key topic in Natural Language Processing (NLP) with the objective of determining sentiment or emotion in a piece of text. Liu [1] proposed a broad review of this topic, covering the approaches and methods used to identify, extract, and classify opinions in unstructured data sources like reviews, blogs, and social media. The paper includes a wealth of material on machine learning and lexical approaches, while outlining many issues in the field, such as sarcasm detection, domain dependency, and opinion spam.

Ensuring an effective and accurate representation of the problem features is central to any text mining task. Ramos [2] emphasises the importance of leveraging the TF-IDF (Term Frequency-Inverse Document Frequency) scheme, as it is a useful method for evaluating each word's importance in a document by relating it to the words in the corpus. TF-IDF is

a statistical measure that helps eliminate noise from common words and increases the contribution of informative words, thus enhancing classification and clustering performance. This was a good method for feature extraction and term weighting.

One of the original applications of machine learning methods to assess sentiment classification was proposed by Pang, Lee, and Vaithyanathan [3]. Their research provided a comparison of Naïve Bayes, Maximum Entropy, and Support Vector Machines, and they presented a demonstration that even simple features such as bag-of-words could differentiate sentiment polarity of movie reviews, and contributed a starting point for models to experimentation with feature selection and model tuning for sentiment analysis. Clustering techniques such as k-means proposed by MacQueen [4] provide fundamental ways to group similar documents for unsupervised learning and exploratory data analysis. When combined with the right distance metrics and feature representations, k-means, despite its simplicity, has proven effective in document clustering, sentiment segmentation, and topic modeling.

PCA has played an essential role in representing text data. Jolliffe and Cadima [5] proposed a modern review of PCA and its extensions, and demonstrate how this method may be useful at reducing high-dimensional data—like word vectors—while retaining the variance and minimising calculations.PCA enhances clustering by converting the original feature space into lower dimensions, which can result in better clustering results and easier-to-understand textual data visualizations. Aggarwal and Zhai [6]proposed key methodologies in text mining, including preprocessing, indexing, pattern discovery, and emerging topics like opinion mining and semantic analysis. Their work is foundational for understanding the architecture of modern NLP systems.

Sebastiani [7] presents an extensive survey of machine learning methods in automated text categorisation, discussing supervised models such as decision trees, SVMs, and k-NN. The paper identifies common challenges such as high dimensionality, sparsity, and class imbalance while exploring solutions like feature selection, threshold tuning, and ensemble learning. Dietterich [8] proposes ensemble techniques, like Bagging, Boosting, and Random Forests, to improve the accuracy and robustness of classification models. Ensemble methods produce new models by combining multiple base learners to help reduce variance and bias, and ensemble techniques can improve interpretability in noisy, sparse feature spaces often associated with text classification tasks, where many classifiers might provide unstable results.

Performance metrics used in classification tasks are systematically analysed by Sokolova and Lapalme [9]. They stress the significance of precision, recall, F1-score, and ROC-AUC and contend that accuracy alone is inadequate, particularly in unbalanced datasets. Their research provides useful recommendations for assessing sentiment analysis and text classification models. Sentiment analysis and text classification are now being applied to human resources and organizational behaviour. Edwards and Edwards [10] show how predictive analytics, especially sentiment analysis of employee feedback, can support HR decisions. Using machine learning and business intelligence tools together, organizations can predict attrition and what drives engagement and plan their workforce more effectively.

Fusco et al. [11] present a novel artificial intelligence framework that combines text sentiment analysis and physiological metrics in real-time to enhance employee well-being. They present a high-performing, multimodal AI framework that combines text sentiment and physical data to generate actionable insights into well-being. The study offers a progressive model for employee wellness in Industry 5.0 with strong measurement models and a useful KPI structure. However, more effort is required to improve the adoption of ethical governance, transparency, and integration with current HR platforms. Although it uses multimodal inputs, their model, which combines wearable data with NLP-derived dissatisfaction scores, is similar to the hybrid architecture of the suggested system.

The literature emphasises how sentiment analysis has developed from basic lexical techniques to complex machine learning models backed by strong feature engineering and assessment techniques. Many contemporary NLP systems are built on techniques like TF-IDF, PCA, k-means clustering, and ensemble learning. Additionally, these techniques are now more applicable outside of academia in fields like HR analytics, where text-based insights are increasingly used to inform data-driven decisions.

3. Proposed Methodology - Sentiment-Driven Productivity

The proposed system is designed to actively monitor and improve workplace happiness, which aims to boost productivity and employee engagement using Machine Learning (ML) and data analytics. In contrast to past methods that are mainly concerned with the simple classification of sentiment polarity, our proposed system puts forward a hybrid framework that fuses the three methodologies of sentiment analysis, employee performance indicators, and employee behavioural profiling through clustering to identify latent patterns in employee satisfaction and productivity. Our key insight is to consolidate the three elements of time-based tracking of sentiment, clustering with TF-IDF, and predictive classification into a single pipeline for a common output, which gives rise to an early warning of group dissatisfaction trends and a type of employee segmentation/targeting. The system collects data from employee surveys, performance indicators, attendance records, and sentiment derived from internal communications. Then, utilising the abilities of

Natural Language Processing (NLP), the system analyses sentiment, applies clustering algorithms to identify employee groups based on their behaviour and performance, and leverages predictive analytics to bring to light the potential causes of dissatisfaction. This feedback text is then processed using Natural Language Processing (NLP) techniques like TF-IDF to prepare it for analysis. The cleaned and structured data is passed into a Machine Learning model trained to classify or predict sentiment and feedback scores. After the model processes the data, the predicted results are sentiment (positive, negative, mixed). Figure 1 illustrates a flow diagram of the proposed algorithm. The Sentiment-Driven Productivity algorithm is defined as follows:

Table 1. Comparison of Machine Learning Techniques for Sentiment Analysis in HR Analytics

Study/Paper	Methodology	Model Type	Features/ Techniques Used	Application Area	Key Contributions
Liu(2012)[1]	Sentiment analysis via NLP	NLP Framework	Subjectivity extraction from text	General Sentiment Analysis	Defined the foundational scope of sentiment analysis
Ramos (2003) [2]	TF-IDF Vectorization	Text Vectorization	Term frequency& inverse document frequency	Text Mining	Popularised TF-IDF as a robust feature extractor
Pangetal. (2002)[3]	Sentiment classification with Naïve Bayes	Probabilistic Classifier	Probabilistic handling of high-dimensional text	Sentiment Analysis	Demonstrated effectiveness of Naïve Bayes in text classification
MacQueen (1967) [4]	Clustering employee feedback	K-Means Clustering	Pattern discovery in text data	HR Feedback Clustering	Introduced K-Means; useful for unsupervised feedback pattern recognition
Jolliffe & Cadima(2016) [5]	Dimensionality reduction	PCA	Feature space reduction & visualization	Text Analytics Visualization	Enhanced data interpretability in high-dimensional feedback spaces
Aggarwal & Zhai(2012)[6]	Unsupervised feedback trend detection	Clustering	Sentiment trends, latent issues	Employee Sentiment Analysis	Used clustering to uncover hidden HR sentiment patterns
Sebastiani (2002)[7]	Supervised text classification	Logistic Regression, SVM, etc.	Feature-based classification models	General Sentiment Classification	Benchmarked classical ML models for sentiment classification
Dietterich (2000)[8]	Ensemble learning	Voting Classifier	Combination of multiple ML models	Robust Sentiment Prediction	Improved accuracy and stability with ensemble approaches
Sokolova & Lapalme(2009) [9]	Model evaluation techniques	Evaluation Metrics	Accuracy, precision, recall, F1-score	ML Model Performance in Sentiment Analysis	Standardised performance assessment metrics for classification models
Edwards & Edwards(2019) [10]	HR-focused text mining	Applied ML in HR	Feedback-driven decision support	HR Analytics and Policy Making	Connected sentiment analysis to HR improvement strategies
Fusco et al. [11	Sentiment Analysis of employees	Applied multimodal AI	Progressive model	Industry decision making	Connected to sentiment analysis

3.1. Create a Dataset and Input Employee Feedback

The dataset is created by compiling written feedback from employees, along with productivity and attendance data. Data was collected from 2007 in various organizations working in various departments. Figure 2 shows the feature distribution of employees across departments. The data collected can be represented as follows:

Let the set of employee feedback be: $Fe = \{fe_1, fe_2, ..., fe_n\}$ where fe_i is the textual feedback from the i^{th} employee.

3.2. Pre-Process Text from the Employee's Feedback

Feedback text is processed using Natural Language Processing (NLP)—including text cleaning, tokenisation, and TF-IDF vectorisation—to convert unstructured text into structured numerical features. Numerical data is normalised. Missing values are replaced by mean values, and data scaling is done. Each feedback fe_i is pre-processed. Tokenisation is done using a mathematical formula:

$$fe_i \rightarrow T_i = \{t_{i1}, t_{i2}, ..., t_{in}\}$$

and Stop-word Removal by

$$T_i' = T_i \setminus \text{StopWords}$$

3.3. Feature Extraction (TF-IDF Vectorisation)

Each cleaned token list T_i' is transformed into a TF-IDF feature vector: $x_i = \text{TF-IDF}(T_i')$, where $x_i \in \mathbb{R}^d$. Set of all such vectors: $X = \{x_1, x_2, ..., x_n\}$.

3.4. Sentiment Classification

Features are used to train Machine learning models like SVM, Naïve Bayes classifier, decision tree and logistic Regression. The logistic regression model gave the highest accuracy for our dataset. Each classifier C maps each vector x_i to a sentiment label: $C: \mathbb{R}^d \rightarrow \{\text{Positive, Neutral, Negative}\}$; $s_i = S(x_i)$.

3.5. NLP-based Clustering (K-Means Algorithm)

The following sub-steps were performed to perform NLP-based Clustering:

3.5.1. Feature Extraction

TF-IDF Vectorisation is applied to transform textual feedback into numerical features.

3.5.2. Clustering

K-Means clustering (k=2) is used to group employee feedback into two clusters. Group the feedback using k-means clustering by minimising $\sum_1 n \min_j || x_i - \mu_j ||^2$, where μ_j is the centroid of cluster j. Each employee i is assigned to a cluster:

$$C_i = \text{argmin}_j || x_i - \mu_j ||^2$$

4. Implementation

The proposed work was implemented using Python.

4.1. Model Training Tools and Environment

Model training and experimentation were conducted in Google Colab, which offers a cloud-based Jupyter notebook environment ideal for developing and testing machine learning models. The platform provides sufficient resources for real-time execution, visualization, and experimentation. The following tools and libraries were used during the model development process.

- Python Standard Libraries such as os, pickle, and csv were used for file management, data serialisation, and utility operations.
- NumPy and Pandas were utilised for numerical computations, data cleaning, and structured data manipulation.
- Matplotlib and Seaborn enabled effective visualization of data distributions, correlations, and evaluation metrics.
- Scikit-learn was the primary machine learning framework used for data pre-processing, such as scaling, splitting, Text vectorisation (TF-IDF), Model training (Logistic Regression, SVM, Decision Tree), and Performance evaluation (accuracy, confusion matrix).

4.2. Experimental Results

Experimental results demonstrated the effectiveness of the proposed system in accurately classifying employee feedback and assessing overall sentiment. The machine learning models were trained and evaluated on pre-processed feedback data, incorporating productivity, attendance, satisfaction rate, and textual feedback metrics. The system successfully achieved its goal of automating employee evaluation through sentiment analysis and classification, contributing to better human resource decision-making. It produces a scalable and data-driven approach to identify employee well-being and performance trends, thereby efficiently reducing manual review efforts and supporting on-time strategic HR planning.

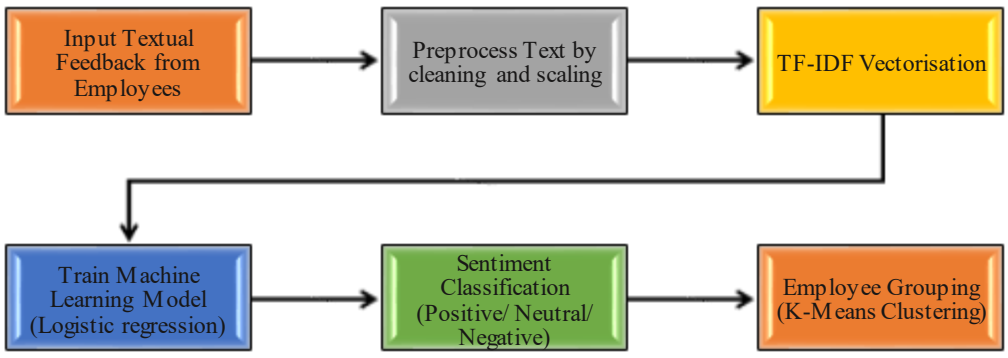

Fig. 1 Work Flow Diagram of Sentiment-Driven Productivity

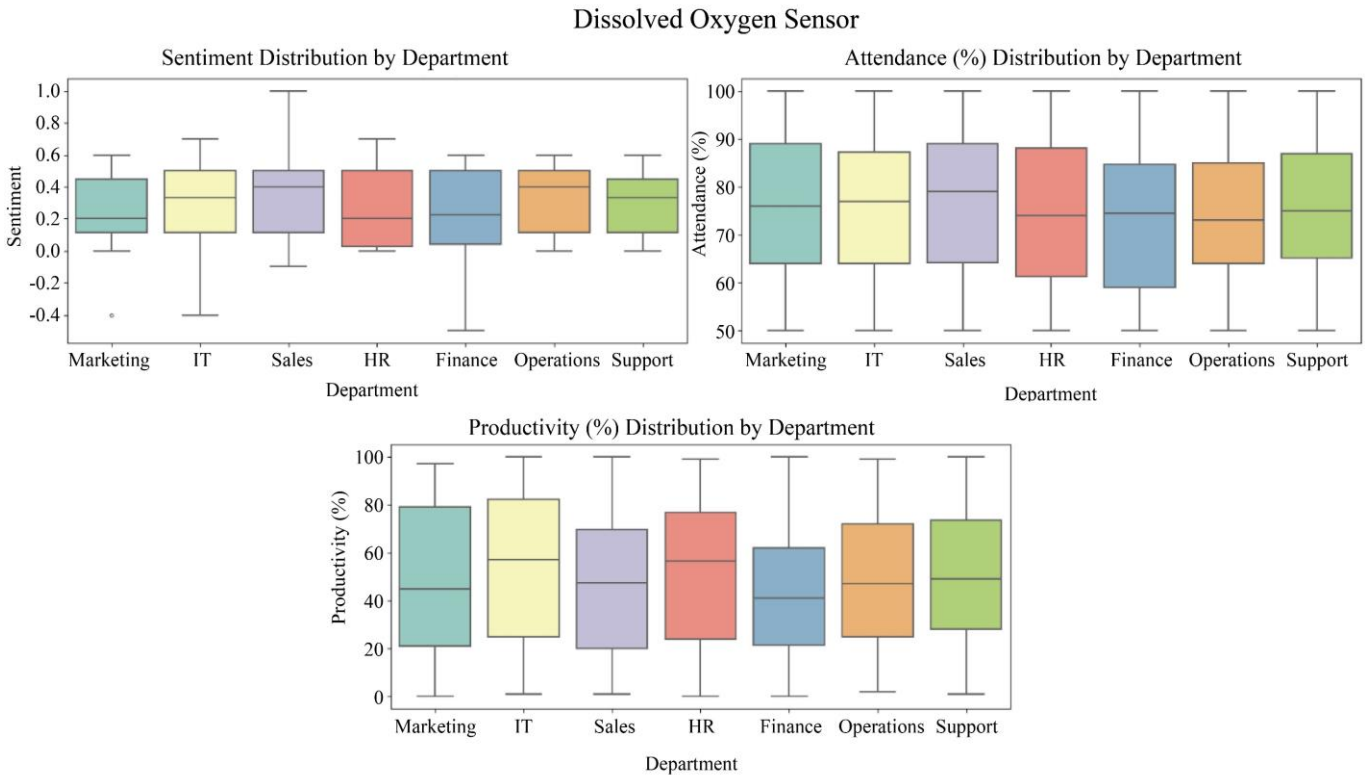

Fig. 2 Feature distribution across various departments

5. Evaluation and Analysis

Accuracy, precision, recall, and F1-score [9] constitute the performance evaluation and comparison of various models. Visualization tools such as confusion matrices and classification reports further confirmed the robustness of the models.

Among these models, the Logistic Regression is found to outperform others with an accuracy of 99.01%, showcasing its effectiveness in sentiment classification tasks. This shows how well employees' happiness is correlated to their attendance and productivity. It maintained a high precision, recall, and F1-score level, indicating a balanced performance in detecting and correctly classifying feedback.

Table 2. Model evaluation metrics

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Naïve Bayes	93.07	93.30	93.07	93.10
SVM	96.04	96.21	96.04	96.07
Decision Tree	91.09	91.35	91.09	91.13
Logistic Regression	99.01	99.05	99.01	99.01
Random Forest	95.54	95.77	95.54	95.55

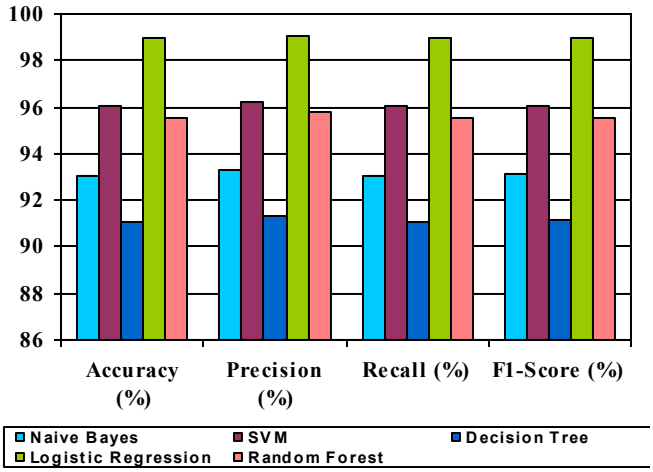

Fig. 3 Performance Comparison Bar Chart for Various Models

Support Vector Machine (SVM) and Random Forest models also performed strongly, achieving accuracies of 96.04% and 95.54%, respectively. While the Naïve Bayes model showed decent performance, it was slightly lower in all metrics compared to the others, though it still remains a viable, lightweight model for large-scale classification. The Decision Tree model had the lowest performance among the selected algorithms, with an accuracy of 91.09%. Confusion matrix of various models is plotted in Fig. 5, Fig. 6, Fig. 7, Fig. 8 and Fig. 9.

The following dataset is used to train employees using the K-Means clustering algorithm. Table 3 shows major features in the dataset, and Figure 10 shows cluster visualization after K-Means clustering, and Figure 11 depicts the results of K-Means clustering. The Silhouette Score is approximately 0.28, which is a low to moderate value. So, for clustering employees, a more efficient algorithm like DBSCAN or hierarchical clustering may work well.

Table 3. Features used to group employees

Column names	Description
Employee ID	A unique identifier assigned to each employee.
NAME	The name of the employee.
Age	The age of the employee in years.
Gender	The gender of the employee (e.g., Male, Female, Other).
Projects Completed	The total number of projects successfully completed by the employee.
Productivity (%)	The productivity level of the employee is expressed as a percentage.
Satisfaction Rate (%)	The employee's job satisfaction is expressed as a percentage.
Department	The department in which the employee works (e.g., Marketing, HR, IT).
Position	The job title or designation of the employee.
Joining Date	The date when the employee joined the company.
Salary	The annual salary of the employee is in numerical value.
Feedbacks	Employees' feedback or comments about their job and work environment.
Sentiment	Sentiment analysis of the feedback (e.g., Positive, Neutral, Negative).
Attendance (%)	The employees' attendance record is expressed as a percentage.

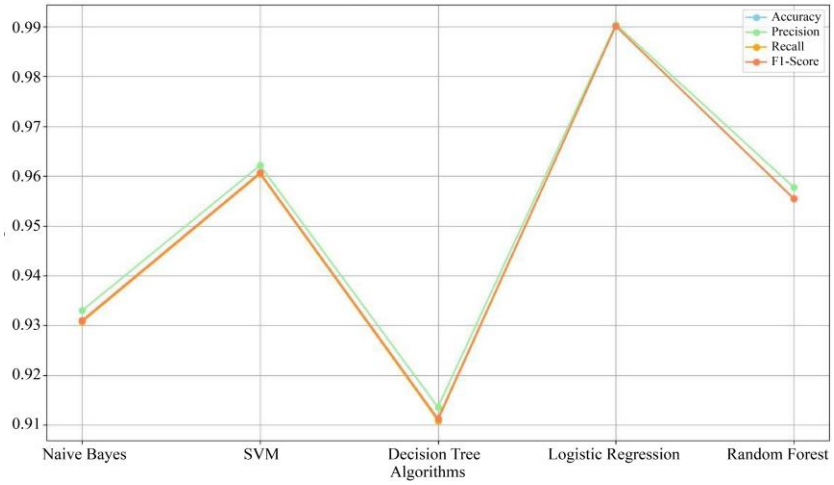

Fig. 4 Line Plot for Algorithm Performance Metrics

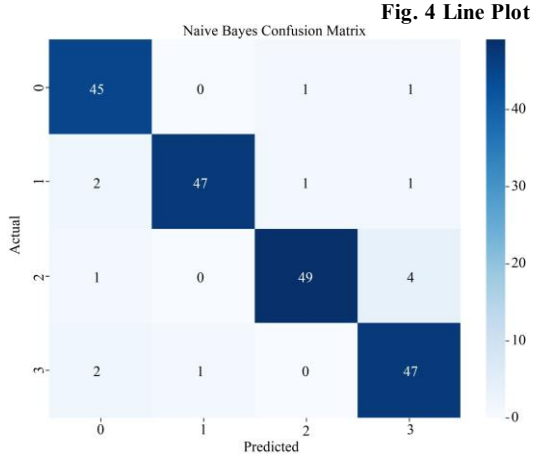

Fig. 5 Confusion Matrix of Naïve Bayes

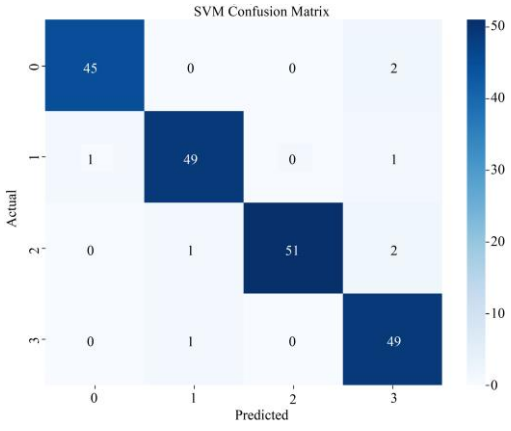

Fig. 6 Confusion Matrix of SVM

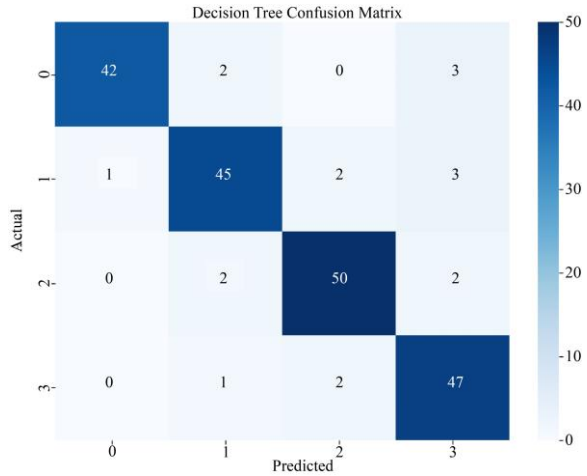

Fig.7 Confusion Matrix of Decision Tree

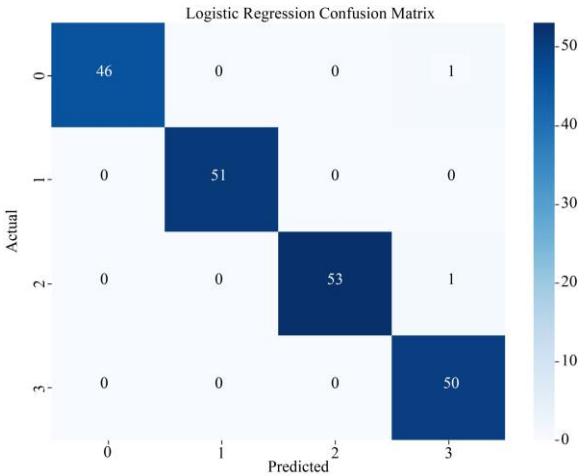

Fig. 8 Confusion Matrix of Logistic Regression

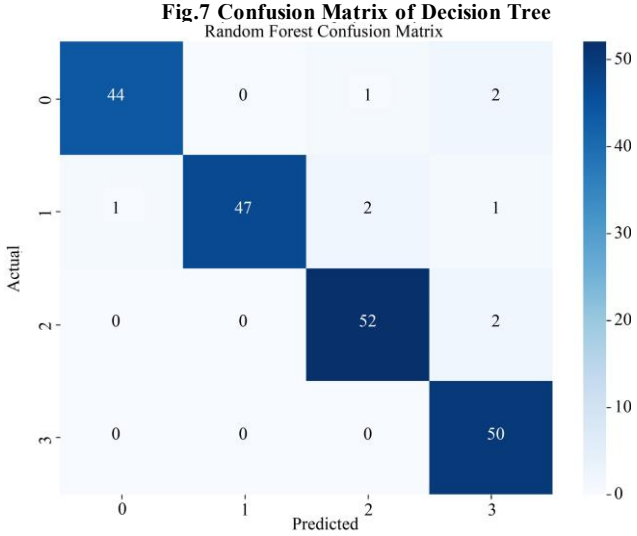

Fig. 9 Confusion Matrix of Random Forest

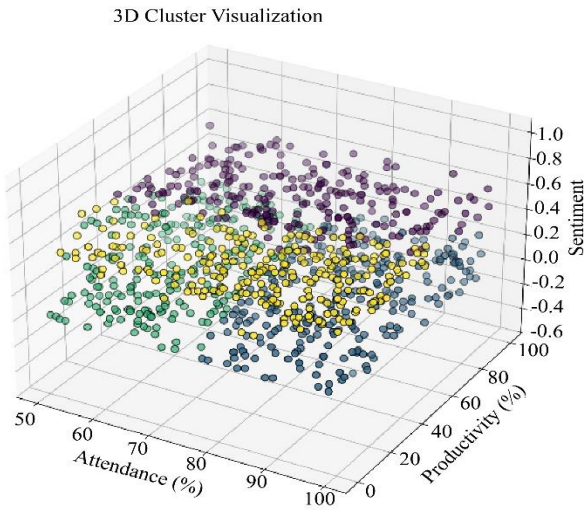

Fig. 10 Cluster Visualization

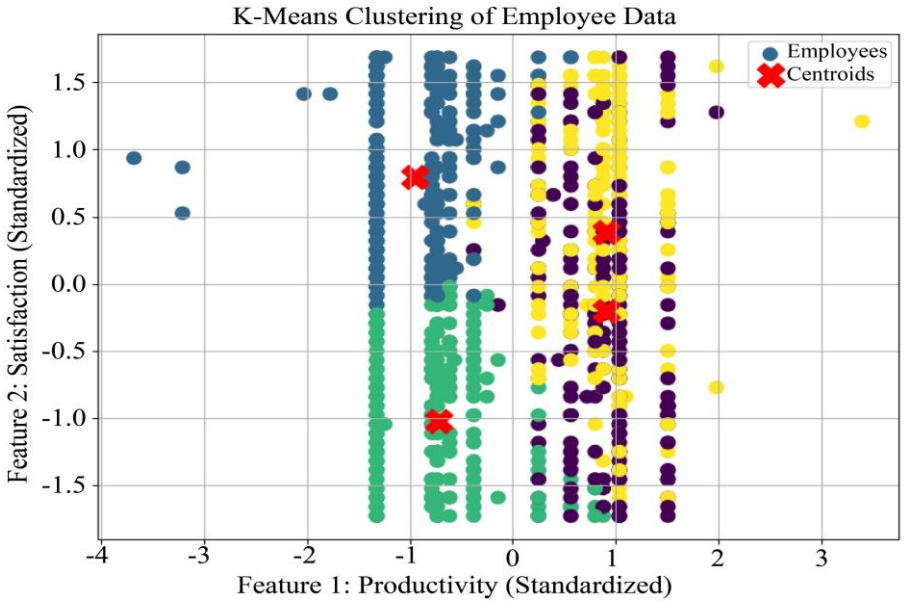

Fig. 11 Result after K-Means clustering

6. Conclusion

The proposed Employee Feedback Analysis System offers an intelligent and automated approach to understanding employee sentiment by utilising machine learning and natural language processing techniques. The novelty of our work lies in the dataset creation and integration of TF-IDF vectorisation and a machine learning algorithm. At its core, the system integrates several supervised learning algorithms, including Random Forest, Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbours (KNN), Naïve Bayes, and Decision Trees to classify feedback into positive or negative categories with high accuracy. These models process the textual feedback to identify and analyse the sentiment and classify the responses based on the emotional tone conveyed by employees. Among the machine learning models used in our work, Logistic Regression is found to outperform others with an accuracy of 99.01%. In addition to classification, the system employs unsupervised learning techniques such as K-Means clustering to group similar feedback entries. But in our work, the Silhouette Score is low to moderate, which can be improved by using other clustering algorithms.

Acknowledgment

We acknowledge the productive contributions of the scientific community in sentiment analysis, NLP, and machine learning, which have informed the direction of this study. Special thanks to the developers and researchers behind the pioneering techniques such as TF-IDF, Naïve Bayes, K-Means clustering, and ensemble learning, whose innovations continue to drive the employee feedback analysis forward. This synthesis would not have been possible without the collective efforts of many researchers over the past five decades. We also thank Prajyoti Niketan College, Pudukad, for their support, guidance, and resources, which enabled the successful completion of this article.

References

[1] Bing Liu, "Sentiment Analysis and Opinion Mining," *Computational Linguistics*, vol. 5, no. 1, pp. 511-513, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[2] Juan Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries," *Proceedings of the First International Conference on Machine Learning*, vol. 242, pp. 133-142, 2003. [Google Scholar]

[3] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan, "Thumbs Up? Sentiment Classification Using Machine Learning Techniques," *arXiv:cs/0205070*, pp. 1-9, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[4] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 5, no. 1, pp. 281-297, 1967. [Google Scholar]

[5] Ian T. Jolliffe, and Jorge Cadima, "Principal Component Analysis: A Review and Recent Developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp. 1-16, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[6] Charu C. Aggarwal, and ChengXiang Zhai, "A Survey of Text Classification Algorithms," *Mining Text Data*, pp. 163-222, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[7] Fabrizio Sebastiani, "Machine Learning in Automated Text Categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[8] Thomas G. Dietterich, "Ensemble Methods in Machine Learning," *Multiple Classifier Systems*, pp. 1-15, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[9] Marina Sokolova, and Guy Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[10] Martin R. Edwards, Kirsten Edwards, and Daisung Jang, *Predictive HR Analytics: Mastering the HR Metric*, Kogan Page Publishers, 2019. [Google Scholar] [Publisher Link]

[11] S. Ranjit Kumar, "Sentiment Analysis on Employee Layoffs Based on Hybrid Feature Extraction and Long Short Term Memory Network," *International Journal of Natural Language Processing*, vol. 2, no. 1, pp. 12-20, 2024. [Publisher Link]