Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.12 n.57, p. 1197-1206. January/February, 2026.

DOI: https://doi.org/10.5935/jetia. v12i57.2741

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

PARTICLE SWARM OPTIMIZATION AND ENSEMBLE METHODS FOR ABSENTEEISM PREDICTION

Baydaa Hadi Saoudi[1], Tabarek Alwan Tuib*[2] and Yaqdhan Mahmood Hussein[3]

[1]Technical Institute-Samawa, Information Technology and Communication (ICT), Al-Furat Al-Awast Technical University,Samawa, Iraq.
[2, 3]Collage of Engineering, Electronics and Communications Engineering, Al Muthanna University, Muthanna, Iraq
[3]Al Ayen University, Nasiriyah 64001, Iraq

[1]https://orcid.org/0000-0002-8198-324, [2]https://orcid.org/0000-0002-1646-0015, [3]https://orcid.org/0000-0001-7153-4120

Email: ins.byd44@atu.edu.iq, *tabarcka@gmail.com, Yaqthanm79@gmail.com.

ARTICLE INFO

Article History
Received: September 25, 2025
Revised: October 20, 2025
Accepted: December 15, 2026
Published: February 28, 2026

Keywords:
Absenteeism,
Particle Swarm Optimization,
Ensemble classification.

ABSTRACT

As it might lead to the loss of productivity and an economy, high absenteeism of employees badly affects organizations. With regard to courier companies, the methods of machine learning (ML) are utilized for the absenteeism of employees. The research uses inductive learning method, using conventional methods of data mining (DM), like data preprocessing, feature selection (FS) by Particle Swarm Optimization (PSO), and data classification by using such ensemble techniques, like Boosting, Voting, Bagging, and Stacking. The dataset with 21 attributes were 740 instances and this was processed to eliminate irrelevant features and transformed for effective analysis. The proposed model was very accurate with Boosting being the most successful ranking at 99.8 accuracy. The research indicates that FS is significant in enhancing accuracy of predictions and it is a useful instrument that managers can use to reduce the effects of absenteeism in organizational performance. In future, it will be possible to incorporate additional datasets, investigate more advanced methods of FS and to incorporate external variables to improve prediction.

I. INTRODUCTION

ML is regarded as a pivotal domain in research, aimed in uncovering or forecasting subjects [1]. Artificial intelligence enables machines to exhibit cognitive functions. Furthermore, artificial intelligence enhances the cognitive capabilities of machines. ML constitutes a domain within the field of artificial intelligence study.[2] Certain researchers assert that intelligence cannot be cultivated without education. [3] All of these methods have been used for categorizing the subjects of dataset. In supervised learning, a training set of examples with corresponding targets is supplied, and algorithms responding properly to all the potential inputs depending on this training set. Supervised learning encompasses classification and regression. [4] The semi-supervised learning methodology can be defined as a variant of supervised learning. In addition, unlabeled data was used for the training task in this approach. [5] Furthermore, DL is a part of ML built on a set of algorithms. [6] There is an urgent need for ML approaches in today’s environment in order to reduce human effort while reducing errors and achieving the greatest degree of automation.[7] Absenteeism is widely seen as a significant signal of inadequate performance inside a business.[8] Absenteeism from work has negative impact on the country's gross domestic product.

Furthermore, the annual expense of absenteeism in nations is approximated to be between 1.2 and 2% of GDP.[9] Absenteeism among blue-collar workers may stem from several factors, including religious and social commitments, inadequate housing, also subpar working circumstances. Organizations employ several strategies to mitigate absenteeism, such as adopting work rotation and enrichment, providing financial incentives and bonuses, and enforcing punctuality regulations. [10] Absenteeism is a pertinent issue in human resource management across many industries and organizations [11], [12] in the context of international human resource management (IHRM). The Micro-HR management can be defined as the sub-functions regarding HR policy and it has two main categories: the first one is related to the management of people and small groups, which involve the selection and recruitment, orientation and training, performance management, in addition to rewards. The other is work organization management and employee voice systems [13], [14]. The issue of absenteeism is a major obstacle which must be tackled by any company and employee.

The employees should have a sense of fulfillment in their employment and the employers should help them have this good experience in every possible way. Absenteeism is a common pattern of non-attendance of some responsibility or duty. The issue of absenteeism in employees is not purely a financial issue; it has a great effect on employee morale as well. Absenteeism in the workplace can be a pointer to low morale among the employees and management [15]. Therefore, one of the important issues to consider when making predictions of human behavior, especially absenteeism, is the means in which large datasets can be shifted or refined. The FS is an initial step in the process of DM. It aims at removing the extraneous and redundant features of the original data [16], [17]. In most cases, companies strive to increase their profits.

On-demand, casual, and temporary employment (which is common in the gig economy) can be used in organizations where employees are involved in simple tasks because of a lesser number of contractual obligations. The importance of employee specialization and job permanence are increased in companies where employees are doing specialized tasks. The importance of skills and knowledge has been defined and the ability to engage in learning continuously has been an urgent need to firms. The paper will be structured as mentioned: the second part clarifies a number of approaches and evaluates the relevant literature, whereas the third part outlines the data analysis, feature extraction, and CV classification techniques, and describes the proposed ML system. Part IV outlines the experimental results as well as the analysis of results. The fifth part compares the proposed model with the existing studies. The conclusion is marked in the end by Section VI.

II. THEORETICAL REFERENCE

Prior study endeavored to employ various data mining techniques or integrate distinct models to address the issue of absence prediction. Martiniano and colleagues. A neuro-fuzzy network was created using a multilayer perceptron design in an artificial neural network, implementing the error backpropagation method. This report offers initial findings from the utilization of the neuro-fuzzy network in forecasting workplace absence [18], [19]. A fuzzy neural network was constructed using error backpropagation algorithm with a multilayer perceptron to forecast absenteeism. Nonong and colleagues A study was undertaken to develop decision tree designed for detecting prevalent traits regarding employees who often exhibit absenteeism from work [20]. In a recent research, Gayathri employed naive Bayesian classifiers, multiclass perceptron, also J48 classifiers for developing classification model for predicting the employee absenteeism over extended or short durations [21].By [22] They employed artificial neural networks in their study.

Furthermore, in a pertinent study [23], researchers advocated for using neural networks (NNs) as well as DL algorithms for predicting employee commitment behaviors within the workplace. Conducting comprehensive research on large amount of data that is collected via businesses is a laborious and resource intensive task. Too much irrelevant information makes it hard to understand and comprehend [24]. As a result, effective filtering and narrowing down of large volumes of data is one of the focal points of human behavior prediction, especially when it comes to absenteeism. One of the most important preprocessing stages in DM is FS, which tries to remove irrelevant and redundant features of the original dataset [25], [26]. Addition of extraneous and redundant data into the model can lead to high expenditures, long time wastage, and reduction in the accuracy [27], [28]. The existing literature focuses on the lack of studies that use DM-based research to predict absenteeism and the lack of emphasis on the crucial aspect of FS.

By [29] noted that no studies aimed at the development of appropriate methods to compare and evaluate the effectiveness of different DM classification methods and the current literature also content highlight the lack of research using DM methods to predict absenteeism, which has led to the description of the absence of emphasis on the critical aspect of FS. Noted that there is no research aimed at developing adequate approaches to compare and evaluate the effectiveness regarding various DM classification methods and specific sets of features [29]. An alternative ML algorithm is applied in [30]. The study uses the linear regression and support vector regression (SVR) for the purpose of coming up with prediction model of absenteeism. Linear regression is performed with the use of all characteristics, and SVR is performed with consideration of two variables, namely, age and season. The results of the two algorithms are used to show that they are effective in predicting occupational absenteeism with high accuracy [31].

The proposed research will search and evaluate suitable ML tools to predict and evaluate workplace absenteeism. Parameters in the information being reviewed are age, education, employment category, day, month and term of service consisting of 125,000 records in total. The Weka tool is used to compare and contrast different algorithms in terms of accuracy, precision, and sensitivity [32]. This paper is an analysis of the absenteeism within courier company located in Brazil. Methods of ML were applied for understanding and prediction. It has already passed through preprocessing and different ML classification methods have been applied and these include ZeroR, J48 tree-based, k-NN, and NB classification methods. The study introduces the algorithms which could predict the absenteeism to accuracy of 92%. In addition, disciplinary failure has come out as one of the most predictive of absenteeism among the first 20 qualities [33].

The purpose of the presented work was discovering significant aspects as well as the best classification algorithms enhancing predicting absenteeism accuracy. We utilized FS methods for improving the efficiency regarding absenteeism prediction. The results of the experiment proved the importance of the selected features. Moreover, bagging has the best performance with accuracy of 92% in the seven prominent methods that apply CFS in FS [34]. Absenteeism is a major threat to businesses in the world. Our experimental group used logistic regression (LR), random forest (RF), XGBoost, and SVM for analyzing the data of 32 blue-collar employees within fast-moving consumer goods organization between 2017 and 2019. XGBoost and RF possessed F1 of 0.6399 and 0.6461, respectively, indicating fair predictive accuracy of absence. However, there are major growth prospects [35].

III. ADOPTED METHODOLOGY

III.1 PROPOSED METHOD

The research will attempt to formulate a predictive model of employee absenteeism in a courier company through inductive learning model. The model employs the conventional DM, which involves the application of Scikit-learn in Python. The model follows 3 stages including data preparation, FS with the help of PSO, and classification using ensemble methods.

It is aimed at precisely predicting absenteeism to help managers make decisions. This model applies preprocessing, PSO, bagging, boosting, voting, and stacking techniques in order to boost the predictivity. The performance related to integrated classification model is evaluated in terms of statistical metrics.

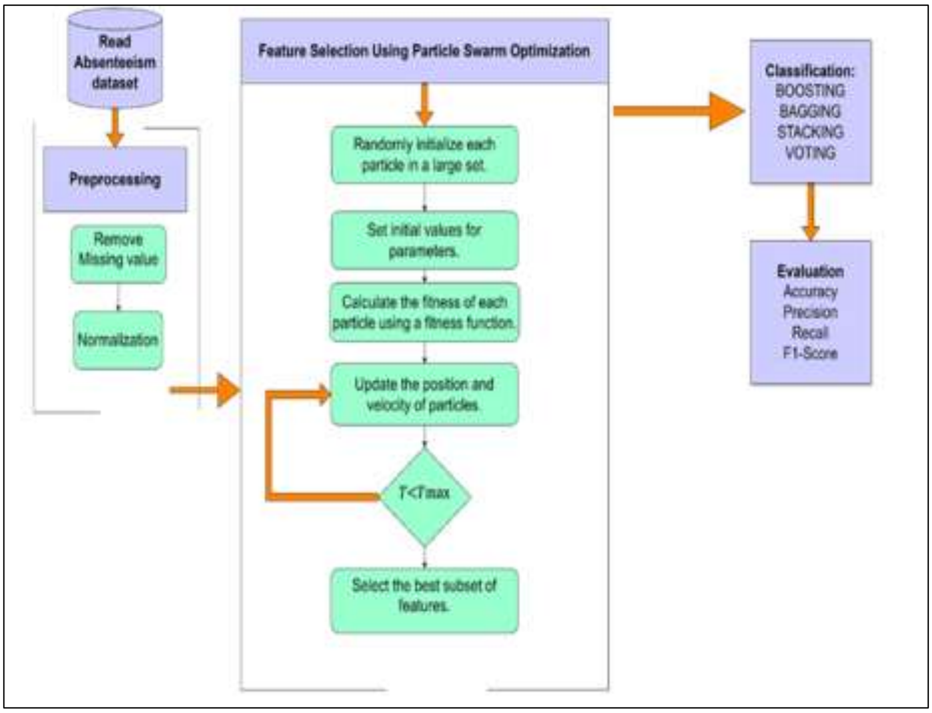

Figure 1: The proposed methodology for resume classification.
Source: Authors, (2026).

III.2 DATA SET

In this paper, we will examine absenteeism in courier company in Brazil using a dataset. The dataset is on Kaggle and records employees' attendance and cause of absenteeism among the employees, and other factors. It assists in determining trends and offers information on factors that have an impact on absenteeism. There are also work-related parameters such as work load and achievement targets. This dataset consists of 740 incidences of more than 21 attributes in three years. The paper proposes ML solutions to successful human resource strategies and employee interaction. The DL techniques, however, are not used because of its small size. Data is broken into training and test data.

III.3 FEATURE SELECTION USING PARTICLE SWARM OPTIMIZATION

III.3.1 Particle Swarm Optimization

In the PSO algorithm, each one of the particles is represented as a vector containing continuous values between 0 and 1. Each element in this vector represents a particular feature in the dataset. These continuous values are converted to binary values (0 or 1) using a binary conversion function. In the case when the value is greater than a specified threshold (0.5), it is assigned to 1 (feature is selected), and if it is less than the threshold, it is assigned to 0 (the feature is not selected).

III.3.2 Fitness Function

The objective function is the criterion by which the quality of each particle (solution) is evaluated. In this code, objective function is designed to evaluate the performance regarding a selected set of features based on two main criteria:

III.3.2.1 Error Rate

A classification model (such as Bagging, Boosting, Stacking and voting) is trained using the selected features from the particle. The error rate is then calculated depending on the model's performance on validation data. The goal is to minimize the error rate as much as possible.

III.3.2.2 Feature Size

The number of selected features is minimized to avoid overfitting and improve model efficiency. The goal is to select the fewest possible features while maintaining classification accuracy

$$\text{fitness} = \text{ErrorRate} * \alpha + \beta * \frac{\text{Number of Selected Features}}{\text{Total Features}} \tag{1}$$

The importance of each criterion is controlled using the parameters α and β, where α represents the weight of the error rate (0.99), prioritizing classification accuracy, while β = 1 - α represents the weight of the number of features (0.01), reducing the impact of the number of features compared to the error rate. The error rate is calculated based on the performance of the classifier, and the number of selected features is the number of features selected from the total number of features

III.4 ENSEMBLE CLASSIFICATION

III.4.1 Bagging

Bagging is implemented with the use of RF algorithm, where random samples are taken with the return from the training data. Multiple independent decision trees are trained on these samples, and their results are then combined utilizing majority voting in classification. The figure 2 illustrates how Bagging works, where random samples of data are taken, sub-models (RF) are trained in parallel on these samples, and then their results are combined to obtain the final prediction

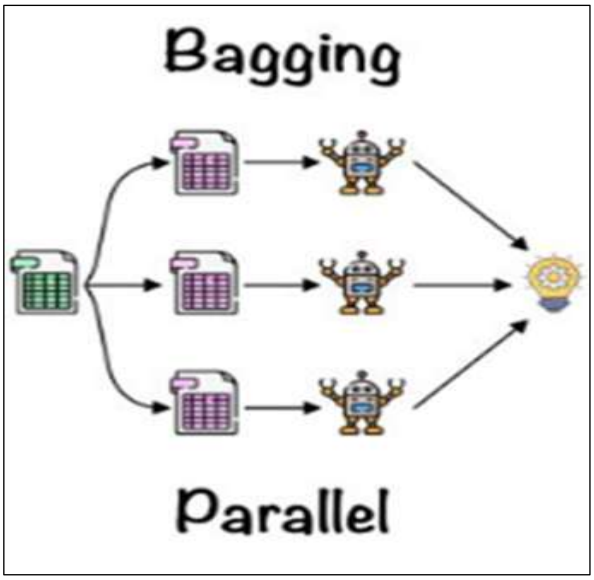

Figure 2: Bagging process.
Source: Authors, (2026).

III.4.2 Boosting

Boosting is based on the RF, and weak models are trained in sequence. The weight assigned to the samples that were misclassified at every step to enhance their performance in the successive models is higher. The outcome of the models is progressively increased in a way that enhances a stronger and more accurate classifier. The figure 3 depicts the Boosting process in the project whereby the models (RF) are trained in series, and all the models enhance the performance regarding other models upon the errors of preceding model and finally, the results are combined to acquire a strong classifier.

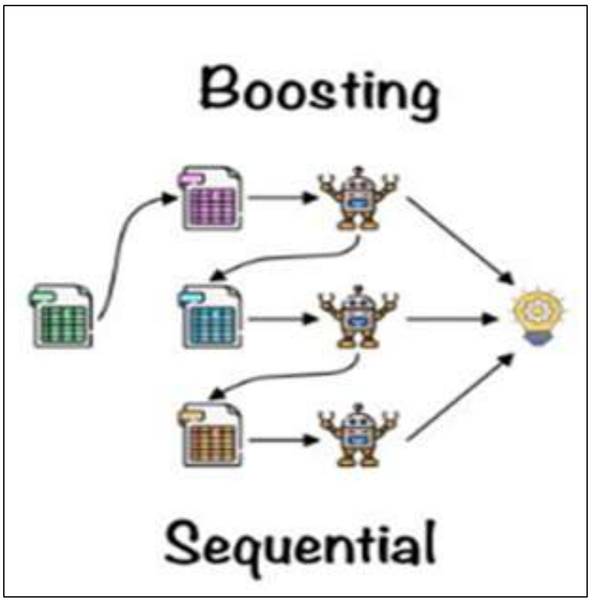

Figure 3: Boosting process.
Source: Authors, (2026).

III.4.3 Voting

The results of three basic models (LR, NB and SVM) were combined by voting in order to enhance the classification accuracy. This is the method by which hard voting is applied to the final decision which is made based on the principle of voting. Both models are applied to the same dataset and the final choice in the models is made based on the majority of the predictions of the models. The figure 4 presents the performance of three simple classifiers (LR, NB, and SVM) on the same dataset. The voting component is used to combine the results of these classifiers and make the final decision based on the majority.

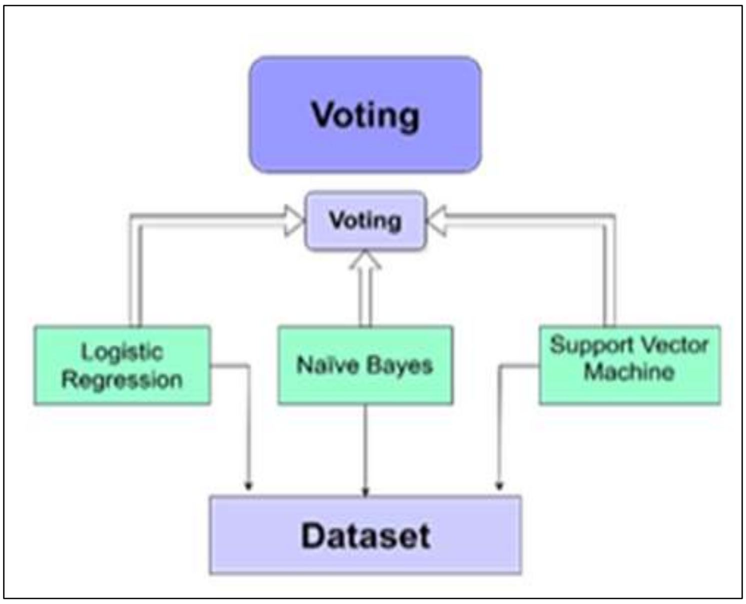

Figure 4: Voting process.
Source: Authors, (2026).

III.4.4 Stacking

The method is dependent on the training of a number of weak base models, including LR, k-NN, and SVM, on training data. The predictions of these models are used as features to train a stronger meta-model, which in this research is Gaussian Naïve Bayes. The aim of this process is combining the strengths related to weak models to obtain a more accurate and stable model. The figure 5 illustrates the two-level stacking process:

Level 0: Weak models (LR, KNN, and SVM) are trained on the training data and generate predictions.
Level 1: The predictions of the weak models are utilized as input to train a stronger model (Gaussian Naïve Bayes), which makes the final prediction.

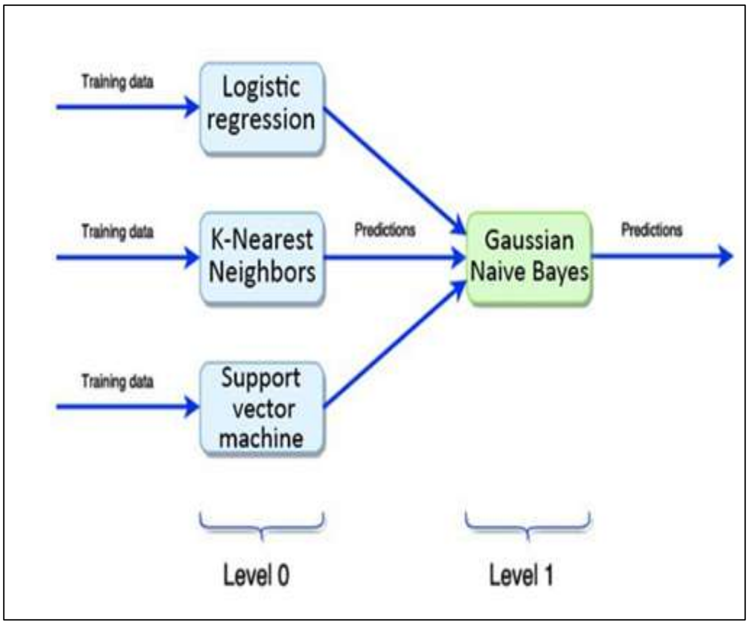

Figure 5: Stacking process.
Source: Authors, (2026).

III.5 ALGORITHM: PSO WITH MODEL (BAGGIN, BOOSTING, STACKING AND VOTING)

Input: Dataset (X, Y), Number of particles (N), Maximum iterations (T), Evaluation criterion (Fitness)

Output: Best feature set, Final model

Step 1. Initialize particles:

- For each particle, randomly initialize a position (feature set) and velocity.

Step 2. Evaluate fitness:

- For each particle, calculate the fitness of the feature set using the evaluation criterion.

Step 3. Update best positions:

- Update the personal best (Pbest) for each particle based on its fitness.

- Update the global best (Gbest) based on the best fitness among all particles.

Step 4. Update velocity and position:

- For each particle, update its velocity using the formula:

Velocity = w * Velocity + c1 * r1 * (Pbest - Position) + c2 * r2 * (Gbest - Position)

- Update the position of the particle based on its new velocity.

Step 5. Repeat steps 2-4:

- Continue for a specified number of iterations (T) or until the stopping criterion is met.

Step 6. Select the best feature set:

- The feature set represented by the global best particle (Gbest) is selected.

Step7. Train Bagging model:

- Use the selected features to train a model on the dataset.

Step 8. Output:

- Return the best feature set and the trained model.

IV. EVALUATION METRICS

In analyzing the results of the proposed models, the use of standard classification metrics was used, like Precision, Accuracy, F1-Score, and Recall. The measures give a detailed insight on the effectiveness of the classifier.

- Accuracy is a general correctness of the model that is specified as the percentage of the correct classification of instances (negative and positive):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

- Precision is the fraction of times positive predictions are made correctly among the instances that are being predicted as positive:

$$precision = \frac{TP}{TP+FP} \tag{3}$$

- Recall (Sensitivity) determines the percentage of correctly recognized positive results relative to the number of all the actual positive cases

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

- F1-Score is the arithmetic mean of Recall and Precision, which provides a balance in the measurement of the two

$$F1 - score \ = \ 2 \times \frac{Recall \times precision}{Recall + precision} \tag{5}$$

These measures provide a reasonable comparison between classifiers, particularly when there is imbalance between classes, since they are not only concerned with the overall accuracy but also with the capability of the model to recognize minor classes correctly.

V. RESULTS AND DISCUSSIONS

In the presented section, the experimental results are given as well as analyzed on the basis of the proposed methodology. The data were split into testing and training sets with training set utilized for developing the models and the testing set being used to test the models. The major aim of the evaluation was to look into the efficiency of PSO in the FS alongside ensemble classification methods Table 1 contains the parameters used to evaluate the effectiveness of various optimization techniques in the field of FS through particle fusion, i.e., Bagging, Boosting, Stacking, and Voting. The widely accepted metrics that were used to determine the model performance included Precision, Accuracy, F1-Score, Recall, and the Area Under the Curve (AUC). Furthermore, the results indicate that PSO dramatically enhanced the models through the removal of redundant and irrelevant features hence improving accuracy and inter-computation efficiency.

Boosting outperformed the other ensemble methods by a significant margin, and the generalization was almost perfect with near-perfect results. Bagging and Stacking also demonstrated high results, which testifies to them being reliable ensemble methods. Although the Voting classifier had slightly lower performance compared to the rest, it still had strong and competitive performance. In general, the findings demonstrate the significance of FS in the process of optimization of models. The proposed framework based on the combination of PSO and ensemble techniques proved to be effective and stable in predicting absenteeism. These performances imply the practical situational applicability of the framework in the management of human resources, and organizations can predict the negative effect of employee absenteeism in their productivity as well as operation proactively.

Table 1: PSO Parameters.

Parameter	Description	Value/Default
N	Number of particles	15
T	Maximum number of iterations	20
w	Inertia weight	0.9
c1	Cognitive acceleration factor	2
c2	Social acceleration factor	2
k	Number of neighbors in KNN	7
thres	Threshold for binary conversion	0.5
alpha	Weight for error rate in objective function	0.99
beta	Weight for feature size in objective function	0.01

Source: Authors, (2026).

IV.1 BOOSTING

The Boosting-based model recorded the best performance with an AUC of 1.0 which is a perfect value in terms of classification in figure 6. The convergence of the PSO fitness values over iterations was found to be very rapid, which proves the efficiency of the algorithm in finding the optimal feature subsets. The performance measures were very high with recall (0.987), precision (0.999), accuracy (0.998), and F1-score (0.993) and Boosting was the highest-performing ensemble method used in this study.

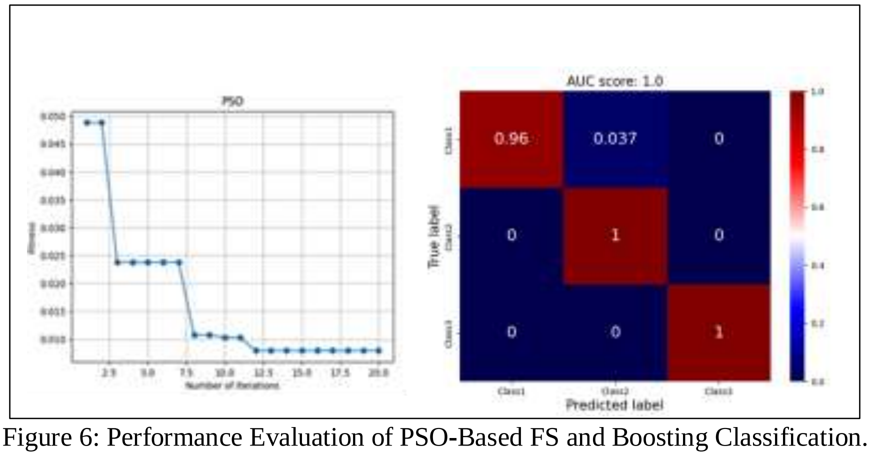

Figure 6: Performance Evaluation of PSO-Based FS and Boosting Classification.
Source: Authors, (2026).

IV.2 BAGGING

The Bagging classifier had an AUC of 0.97 which indicates high classification performance in figure 7. Class 1 had a low level of misclassification (13% to Class 2), whereas Class 2 and Class 3 were classified correctly and 100 per cent. The PSO optimization curve was also characterized by a steady decrease in the values of fitness as the iterations increased maintaining a minimum value, which is indicative of a high degree of convergence. Altogether, Bagging has obtained 0.956 recall, 0.997 precision, 0.993 accuracy, and 0.975 F1-score, which is strong evidence of its strength.

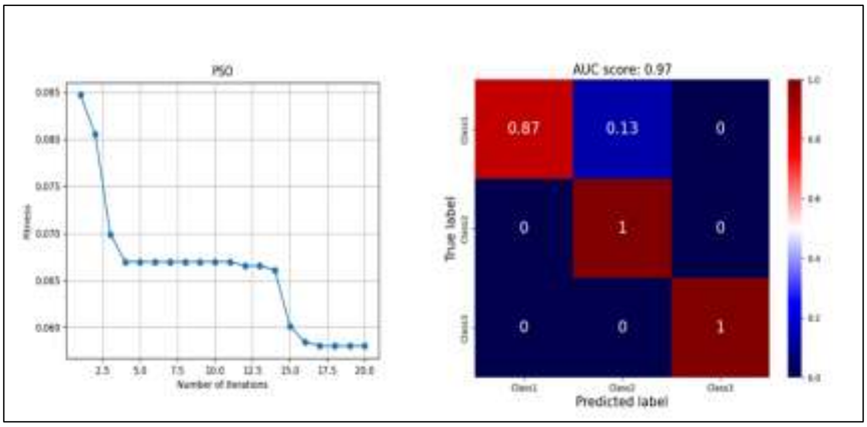

Figure 7: Performance Evaluation of PSO-Based FS and Bagging Classification.
SOURCE: AUTHORS, (2026).

IV.3 STACKING

The Stacking method used weak base models (LR, K-Nearest Neighbors, and SVM), which were in turn followed by a Gaussian Naive Bayes meta-model to fuse their results. The results obtained in this two-level architecture were very reliable with an AUC score of 0.98 in figuer 8 . The technique was able to detect the correct classification of 96 percent of the total data cases which validates that the combination of the multiple learners is effective. The precision, recall, and F1-score of 0.967 as well as the accuracy of 0.990 were the performance metrics used to demonstrate the stability and strength of the stacked model.

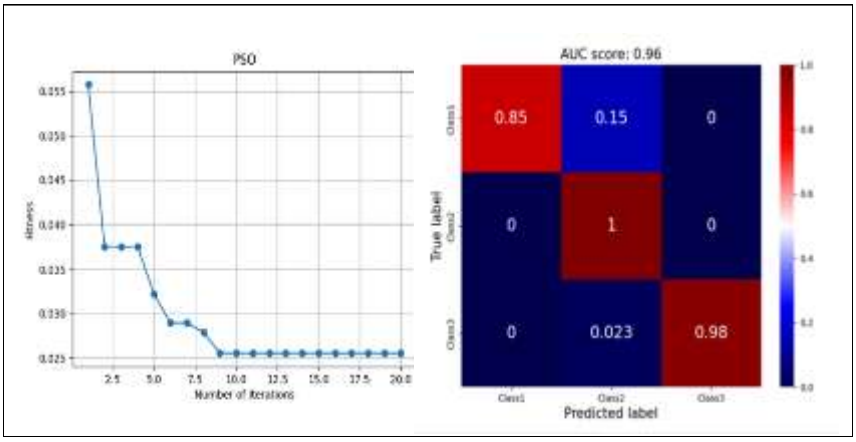

Figure 8: Performance Evaluation of PSO-Based FS and Stacking Classification.
Source: Authors, (2026).

IV.4 VOTING

The Voting classifier used hard voting to combine LR, SVM and NB to come up with the final decision. The approach had an AUC value of 0.981 and was able to classify 98% of the data in figuer 9. Although Voting was less successful in performance than Boosting, it gave precise results with a high recall (0.916), F1-score (0.951), and accuracy (0.982). These findings affirm Voting is a competitive ensemble method as it is used in predicting absenteeism.

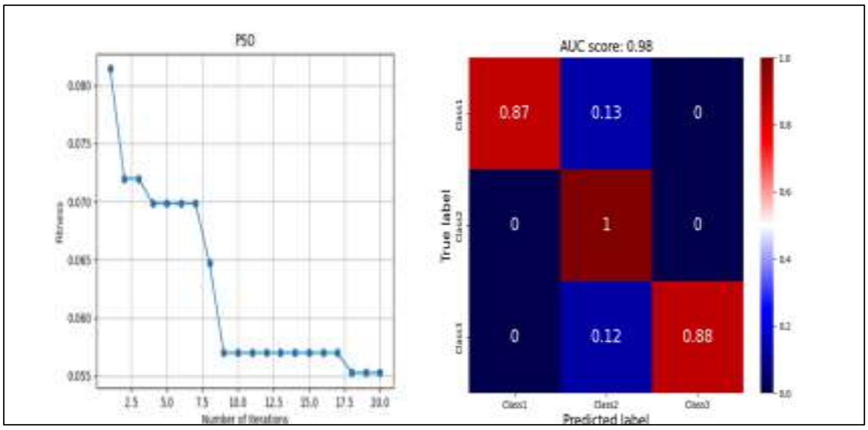

Figure 9: Performance Evaluation of PSO-Based FS and voting.
Source: Authors, (2026).

IV.5 COMPARISON OF MODEL PERFORMANCE METRICS

The performance of the four ensemble techniques Boosting and Bagging Stacking and Voting compares remarkably in table 2. Boosting was the most effective of them with the highest scores in all metrics with recall of 0.987, precision of 0.999, AUC score of 1.0, F1-score of 0.993, and an accuracy of 0.998. Close behind is bagging, which has good results with 0.997 precision, 0.956 recall, 0.975 F1-score, 0.970 AUC score and 0.993 accuracy. Stacking approach was also found to be reliable where the precision is 0.996, recall 0.943, F1-score 0.967, AUC score 0.956 and accuracy is 0.990. Lastly, Voting classifier although slightly lower than the rest, also had competitive scores, showing recall of 0.916, precision of 0.993, AUC score of 0.981, F1-score of 0.951, and its accuracy was 0.982.

These findings affirm that Boosting is the best ensemble technique of predicting absenteeism and Bagging and Stacking have almost similar performance with slight sacrifices in recall. Voting, even though it is slightly lower in performance than the rest, nevertheless, provides good and consistent classification performance.

Table 2: result of evaluation criteria for classification methods.

Method	Precision	recall	f1-score	AUC Score	accuracy
Boosting	0.999	0.987	0.993	1	0.998
Bagging	0.997	0.956	0.975	0.970	0.993
Stacking	0.996	0.943	0.967	0.956	0.990
Voting	0.993	0.916	0.951	0.981	0.982

Source: Authors, (2026).

VI. CONCLUSIONS AND FUTURE WORK

This study proposes a predictive model for employee absenteeism in a courier organization, utilizing PSO for FS combined with ensemble techniques such as bagging, boosting, Stacking, and Voting classification. The methodology involves a three-step workflow: data preprocessing, FS, and ensemble classification using Scikit-learn in Python. Results indicate that PSO effectively removes redundant features and boosts classifier performance, with Boosting achieving the highest accuracy. The model can be used as a decision-support tool by managers and HR specialists in order to control the impacts of absenteeism on productivity through the better management of the workforce. Future studies may include different datasets of various industries, improved FS, incorporation of external variables such as employee satisfaction surveys and use in the real-time decision support systems to improve dynamism in monitoring and predicting absenteeism.

VII. AUTHOR'S CONTRIBUTION

Conceptualization: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Methodology: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Investigation: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Discussion of results: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Writing – Original Draft: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Writing – Review and Editing: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Resources: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Supervision: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.
Approval of the final text: Baydaa Hadi Saoudi, Tabarek Alwan Tuib and Yaqdhan Mahmood Hussein.

VIII. REFERENCES

[1] M. A. A. Albadr, A. S. Tiun, N. S. Alseelawi, and R. M. Al-Saedi, "Speech emotion recognition using optimized genetic algorithm-extreme learning machine," Multimedia Tools and Applications, vol. 81, no. 17, pp. 23963–23989, 2022.

[2] A. A. Barreda-Luna, C. J. L. Bustamante, and F. J. Sánchez, "Spatial models and neural networks for identifying sustainable transportation projects: A case study in Querétaro, an intermediate Mexican city," Sustainability, vol. 14, no. 13, p. 7796, 2022.

[3] M. Fatima and M. Pasha, "Survey of machine learning algorithms for disease diagnostic," Journal of Intelligent Learning Systems and Applications, vol. 9, no. 1, pp. 1–16, 2017.

[4] T. A. Tuib, B. H. Saoudi, Y. M. Hussein, T. H. Mandeel, and F. T. Al-Dhief, "Convolutional neural network with binary mothflame optimization for emotion detection in electroencephalogram," IAES International Journal of Artificial Intelligence, vol. 13,no. 1, pp. 1172–1178, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp1172-1178.

[5] X. Yang, Z. Ding, and T. Liu, "A survey on deep semi-supervised learning," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 9, pp. 8934–8954, 2022.

[6] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," Computer Science Review, vol. 40, p. 100379, 2021.

[7] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," Artificial Intelligence in Agriculture, vol. 2, pp. 1–12, 2019.

[8] A. Cohen and R. Golan, "Predicting absenteeism and turnover intentions by past absenteeism and work attitudes: An empirical examination of female employees in long term nursing care facilities," Career Development International, vol. 12, no. 5, pp. 416–432, 2007.

[9] P. R. de Jong, Sickness, Disability and Work: Breaking the Barriers—A Synthesis of Findings across OECD Countries. Wiley Online Library, 2011.

[10] M. C. Sturman, Multiple Approaches to Absenteeism Analysis. Ithaca, NY: Cornell University, 1996.

[11] P. Bycio, “Job performance and absenteeism: A review and meta-analysis,” Human Relations, vol. 45, no. 2, pp. 193–220, 1992.

[12] D. A. Harrison and J. J. Martocchio, “Time for absenteeism: A 20-year review of origins, offshoots, and outcomes,” Journal of Management, vol. 24, no. 3, pp. 305–350, 1998.

[13] Maria Maqsood, Saima Anwar Lashari, Murtaja Ali Saare, Sari Ali Sari, Yaqdhan Mahmood Hussein and Hatem Oday Hatem, "Minimization Response Time Task scheduling Algorithm", et al 2019 IOP Conf. Ser.: Mater. Sci. Eng. 705 012008, DOI 10.1088/1757-899X/705/1/012008.

[14] C. Murthy, “Employee absenteeism: Its impact on productivity with references to Sukjit Scratch Mills at Nizamabad,” International Journal of Management Research, 2012.

[15] A. M. De Silva, et al., “Feature selection,” in Grammar-Based Feature Generation for Time-Series Prediction, Springer, 2015, pp. 13–24.

[16] A.A. Abdulbari, Z. Zakaria, S.K. Abdul Rahim, Y.M. Hussein, M.M. Jawad, A.M. Hamzah, Design and development broadband monopole antenna for in-doorapplication, TELKOMNIKA 18 (1) (2020) 51, https://doi.org/10.12928/telkomnika.v18i1.13171. Feb.

[17] S. N. Mishra, D. R. Lama, and Y. Pal, “Human resource predictive analytics (HRPA) for HR management in organizations,” International Journal of Scientific & Technology Research, vol. 5, no. 5, pp. 33–35, 2016.

[18] N. Jain and Maitri, “Big data and predictive analytics: A facilitator for talent management,” in Data Science Landscape: Towards Research Standards and Protocols. Springer, 2018, pp. 199–204.

[19] A. Martiniano, et al., “Application of a neuro fuzzy network in prediction of absenteeism at work,” in 7th Iberian Conference on Information Systems and Technologies (CISTI 2012), IEEE, 2012.

[20] N. N. Qomariyah and Y. G. Sucahyo, “Employees’ attendance patterns prediction using classification algorithm: Case study of a private company in Indonesia,” Int’l Journal of Computing, Communications & Instrumentation Engg., vol. 1, no. 1, pp. 1–7, 2014.

[21] T. Gayathri, “Data mining of absentee data to increase productivity,” International Journal of Engineering and Techniques, vol. 4, no. 3, pp. 478–480, 2018.

[22] R. P. Ferreira, et al., “Artificial neural networks and their application in the prediction of absenteeism at work,” International Journal of Recent Scientific Research, vol. 9, no. 1, pp. 2332–2334, 2018.

[23] S. A. Ali Shah, et al., “An enhanced deep neural network for predicting workplace absenteeism,” Complexity, vol. 2020, Article ID 5843932, 2020.

[24] E. L. de Oliveira, et al., “Absenteeism prediction in call centers using machine learning algorithms,” in New Knowledge in Information Systems and Technologies: Volume 1, Springer, 2019, pp. 1–12.

[25] Z. Wahid, et al., “Predicting absenteeism at work using tree-based learners,” in Proc. 3rd Int. Conf. Machine Learning and Soft Computing, 2019.

[26] R. Mythily and D. Mavaluru, “An efficient feature selection algorithm for health care data analysis,” Bulletin of Electrical Engineering and Informatics, vol. 9, no. 3, pp. 877–885, 2020.

[27] R. Kamala and R. J. Thangaiah, “An improved hybrid feature selection method for huge dimensional datasets,” IAES International Journal of Artificial Intelligence, vol. 8, no. 1, pp. 77–85, 2019.

[28] Alaa Abd Ali Hadi, Essam Hamoodi Ahmed, Baydaa Hadi Saoudi, Yaqdhan Mahmood Hussein, Tabarek Alwan Tuib, "Design zigzag edge of S shape slot antenna by using SIW technology for 5G application", TELKOMNIKA Telecommunication Computing Electronics and Control, ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v21i6.23788.

[29] K. Dogruyol and B. Sekeroglu, “Absenteeism prediction: A comparative study using machine learning models,” in Int. Conf. Theory and Application of Soft Computing, Computing with Words and Perceptions, Springer, 2019.

[30] K. Tewari, S. Vandita, and S. Jain, “Predictive analysis of absenteeism in MNCs using machine learning algorithms,” in Proc. ICRIC 2019: Recent Innovations in Computing, Springer, 2020.

[31] A. Rista, J. Ajdari, and X. Zenuni, “Predicting and analyzing absenteeism at workplace using machine learning algorithms,” in 43rd Int. Convention on Information, Communication and Electronic Technology (MIPRO), IEEE, 2020.

[32] M. Skorikov, et al., “Prediction of absenteeism at work using data mining techniques,” in 5th Int. Conf. on Information Technology Research (ICITR), IEEE, 2020.

[33] A. Al-Rasheed, “Identification of important features and data mining classification techniques in predicting employee absenteeism at work,” International Journal of Electrical & Computer Engineering, vol. 11, no. 5, pp. 4083–4092, 2021.

[34] J. Rajath and D. Pandita, “Machine learning as a data science tool to predict absenteeism for factory workers,” in 2022 Int. Conf. on Data Analytics for Business and Industry (ICDABI), IEEE, 2022.

[35] F. Pedregosa, et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.