Received: August 6, 2023. Revised: October 11, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

A Long Short-Term Memory with Recurrent Neural Network and Brownian Motion Butterfly Optimization for Employee Attrition Prediction

Sudhamathy Ganapathisamy[1]* **Valliammal Narayan[1]**

[1]*Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore-641043, India*
* Corresponding author's Email: sudhamathy_cs@avinuty.ac.in

Abstract: Machine learning is a popular technology that continuously supports organizations in their corporate strategies, managerial functions, and team building. There are many areas in which organizations can adopt technologies that will support decision-making. Human resources (HR) have received more attention in recent years as a result of the fact that skilled employees are a significant growth factor and a genuine competitive advantage for the business. Machine learning started to be employed in HR management to help with employee-related decisions after first being introduced to the departments of sales and marketing. The goal is to support decisions that are based on objective analysis of data rather than on subjective factors. In this research, a novel long short-term memory with recurrent neural network (LSTM-RNN) method is implemented for the prediction and classification of employee attrition. To avoid the overfitting issue and make further prediction analysis based on the suitable objective of feature subset selection, the feature selection algorithm termed brownian motion based butterfly optimization algorithm is proposed. The presented model is validated on the international business machines corporation (IBM) HR dataset which consists of 35 features of employees like business travel, age, education, department, daily rate, distance from home, etc. On the dataset of IBM HR, the proposed model achieved accuracy, recall, F score, and precision of 96.68%, 96.62%, 96.62%, 96.64% respectively. The outcomes proved that the proposed method provides better results in the classification of employee attrition.

Keywords: Brownian motion-butterfly optimization algorithm, Employee attrition, Human resources, Long short-term memory, Machine learning, Recurrent neural network.

1. Introduction

The majority of corporate officers are majorly worried about establishing and maintaining a sustainable, effective, and quality workforce because this aspect of business typically strives to be a significant factor in the overall prosperity of the company. All organizations will possibly experience employee attrition. The number of employees gradually leaving the organization due to various factors is called employee attrition. By evaluating retention risk and predicting the possibility of resignations, retention strategies can be established or updated, helping to reduce the high hiring cost and training new employees [1, 2]. The human resource departments of the industries find it challenging to anticipate when and if an employee will leave the organization. By using classification techniques in machine learning (ML), companies can anticipate employee attrition by analyzing employee data. Hence, companies can take effective steps based on predictions to minimize losses [3]. HR functions such as recruitment, training and development, retention, collaboration, and compensation are specific HR functions that are crucial in organizations. Employee attrition analysis in regard to HR analytics has become increasingly popular in the corporate world [4]. One of the most precious assets in today's business sectors is the employee. The effects on the company's competitive edge can make employee attrition a major concern. An organization may encounter expenses as a result of employee attrition. Employee attrition has an impact on the organization's work success, staff morale, misplaced expertise, and the life of human resources (HR) [5].

According to SCIKEY's 2020 talent technology outlook research, the attrition rate is above 22%, as reported in the economic times in December 2019. Financial express analysis from February 2020 states that the IT and BPM sectors contributed about 7.9% of India's GDP in 2019 [6]. By employing supervised machine learning methods, the possibility of employee attrition can be predicted. The data of former employees and current employees can be utilized in designing the prediction model. Various classification techniques like logistic regression, random forest, Naive Bayes, support vector machine methods, and many more can be used for creating prediction and classification models [7]. Towers Watson, a global provider of professional services, claims that attrition in India is remarkably higher at 14% compared to other countries worldwide (11.20%) and in the Asia pacific region at 13.81% [8]. High attrition rates end up being extremely expensive for the business because it spends time, money, and resources training employees to get them ready for a specific company [9]. The prediction techniques will benefit the employees by alerting them to impending attrition and enabling them to make wise decisions and take appropriate action. The outcomes of the prediction will even help the organization to identify employees with higher attrition rates and focus on them for survival [10].

The major contributions to this work are listed below:

- The employee attrition rate is predicted and represented as a time series in which the model can predict the values based on former and current employees. In this work, the IBM HR dataset is considered that consist employee feature set. RNN is used to perform a backpropagation mechanism on LSTM memory cells for the recursive loss reduction and to improve the prediction of attrition rate with less error.
- An LSTM classifier is used to predict the attrition rate categories to achieve accuracy, recall, precision, and F-score. The data is pre-processed and eliminated unwanted features from the dataset. The LSTM classification makes the model robust which is an advantage.
- A brownian-based butterfly optimization algorithm is used for feature selection to select features that contribute to the attrition process. An error value is predicted based on the activation function and sigmoid function.

The rest of the paper is arranged as follows: Section 2 offers the literature survey accomplished in the prediction of employee attrition. Detailed information about the proposed LSTM-based RNN method is given in section 3. The outcomes of the proposed method are provided in section 4. Whereas the comparative analysis is shown in section 5, and section 6 describes the conclusion.

2. Literature review

Fallucchi [11] developed a prediction model to obtain the predicted rate of employee attrition by analyzing the factors that may influence employees in the company. The proposed model was trained and tested on the IBM Analytics dataset, which consists of 35 features and 1500 samples. The proposed framework made use of a Gaussian Naive Bayes classifier to obtain better results. From the overall observations, it is shown that the best recall rate is obtained, which estimates the capability of the classifier to determine the positive and negative instances. The Gaussian Naïve Bayes approach failed to consider employees' intentions over attrition.

Najafi-Zangeneh [12] developed an attrition prediction framework with a three-staged model. The three stages include pre-processing, processing, and post-processing. This model was implemented on the IBM HR dataset and used a logistic regression approach to make use of features and show the importance of those features in the dataset. The dataset consists of several features, out of which one feature selection method known as "max-out" was presented for minimizing the dimensions in the preprocessing stage. The F1 score was improved because of the chosen feature selection method. The logistic regression classifier failed to predict employee attrition with relevant features.

Jain [13] developed an employee churn prediction and retention scheme (ECPR) to predict the turnover of workers, which was an ML algorithm combined with the multi-attribute decision-making scheme (MADM). In this approach, the employees were grouped into different categories by using an accomplishment-based employee importance model (AEIM). The accomplishments of employees were assigned relative weights using an improved version of the entropy weight method (IEWM). Following this, to predict the employee churn in a class-wise manner, a CatBoost algorithm was applied. The proposed method was tested on the dataset of the human resources information system (HRIS) and compared with the existing machine learning algorithms. This approach achieved the highest accuracy, precision, recall, and MCC. However, this developed model not considered employee categorization and category-wise retention strategy.

Ali [14] proposed a prediction system for classifying employee turnover by using NIOPCA (new intensive optimized principle component analysis) and RFC (random forest classifier). The RFC was used for the classification tasks, and NIOPCA was used for feature selection. To decompose the main cause of employee turnover, information mining techniques were used with the random forest classification method. Compared to the existing methods, the performance of this approach achieved high accuracy, precision, F1 score, and recall. The relationship between demographic variables was not described in this work, which is very important to understand the concept of turnover. However, this developed model did not choose influential features.

Cai [15] developed a framework to represent employee turnover with a graph embedding method. The dynamic bipartite graph embedding (DBGE) gives the vector representation of employee relations with the company, including temporary information from the records related to the employee's work. Two approaches, namely the horary random walk (HRW) and the skip-gram model were implemented to produce the vertex sequence in a sequential order and to represent that vertex in low dimensions. To implement this framework, a real dataset was collected from China’s biggest online professional social network. This proposed approach achieved high performance with the graph embedding method compared to the experiments done on Amazon and Taobao datasets that were publicly available. However, this model enhances the training time and capture complex temporal dependencies.

Kakad [16] developed an XGBoost machine learning model to predict employee attrition. This approach is made to address the issues of financial loss that cause to replacement of the trained employee. To help entities take the necessary steps for employee retention or succession in due time with high prediction accuracy. To improve the economic stability of the companies by predicting employee attrition, this approach was made. XGBoost model achieved high accuracy and low memory utilization in prediction tasks. However, this model contains numerous redundant and irrelevant data.

Pratt [17] developed a machine learning algorithm to predict the chances of employees leaving the company. In this paper, different machine learning algorithms were compared and implemented on a real dataset with 1469 samples to find the best algorithm for predicting employee attrition. The considered dataset consists of information regarding former and current employees. The obtained results have shown that the implementation of the random forest algorithm achieved high prediction accuracy. Employee turnover was not considered in the chosen feature selection method, which was a limitation of this work. Also, this model unable to identify real factors and more complex to interpret

Nested ensemble learning techniques (NELT) have been used by Muneera Saad Alshiddy and Bader Nasser Aljaber [18] to show employee attrition prediction. This study looks at how ensemble learning approaches can improve the hiring process by foreseeing staff churn. This study used the IBM HR analytics employee attrition & performance dataset and a two-layer nested ensemble model. As a benchmark for comparison, the efficiency of this framework was contrasted with that of the random forest (RF) algorithm. By employing ensemble models as the basic algorithms in the present work, NELT were explored. Still, this investigation did not go into any further depth to discuss this technique's specifics.

Machine and deep learning models have been used by Samer M. Arqawi [19] to predict employee attrition with a significant amount of accuracy. This investigation aims to assist the human resources department by providing them with the necessary information regarding the likely choice for any employee who would be leaving the company. The suggested model determines whether there is a latent risk of staff attrition. The present research was capable of to provide comprehensive answers to all queries regarding the use of a model that accurately detects staff attrition. Nine machine learning models and one deep learning model were the only ones used in the current investigation. There were numerous machine models that this study could possibly test.

Muslim Lhaksmana Kemas and Sindi Fatika Sari [20] used feature selection with information gain (IG) and random forest (RF) classification to forecast employee’s attrition. To determine which feature selection technique yields the highest performance, this study applies random forest while evaluating information gain, select K highest, and recursive feature elimination. The use of the previous methods beats earlier research in terms of precision, recall, accuracy, and f1 scores. Every time a different feature was applied, the accuracy fluctuated between rising and declining. This does not imply that the study was a failure because, if accuracy rises, the quantity of features was the right amount, and if accuracy declines, the number of features was not the right number.

Future research will focus on greater detail on these new findings including feature selection

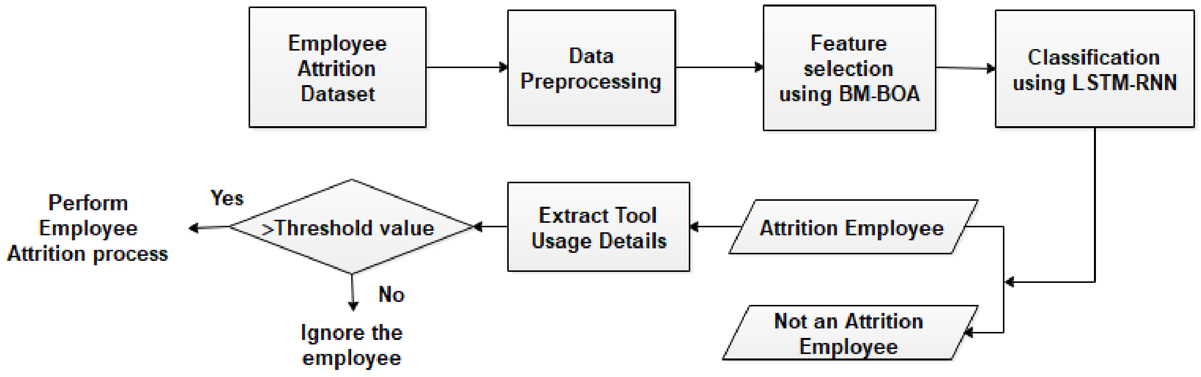
Employee Attrition Dataset
Data Preprocessing
Feature selection using BM-BOA
Classification using LSTM-RNN
Perform Employee Attrition process
Yes
>Threshold value
No
Ignore the employee
Extract Tool Usage Details
Attrition Employee
Not an Attrition Employee

Figure. 1 Architecture of the proposed LSTM-RNN model

techniques, which are an important addition to the data already available. The relevant features on the selection of attributes for the prediction of attrition are not taken into account by existing methods like Naïve Bayes, logistic regression, AEIM, and XGBOOST. This results in an error that recurs through the prediction as loss and cannot be negligible in larger predictions. As a result, these limitations can be overcome by incorporating brownian motion into the butterfly optimization algorithm (BM-BOA) for effective feature subset selection. The RNN's recurrent nature acts as a processing unit in LSTM which results in less error rate of employee attrition prediction.

3. Methodology

The framework of the proposed work is a combination of feature selection with brownian motion based butterfly optimization algorithm (BM-BOA) and long short-term memory (LSTM) classification to predict and classify employee attrition and it is illustrated in Fig. 1. The time characteristics will be taken into account by the LSTM layers during classification, which will be neglected during the feature extraction stage. To optimize the computational resources and make more efficient employee attrition classification, LSTM incorporated in RNN is used. LSTM is combined with RNN which can possess long short memory as well as long-term dependency with the backpropagation nature of RNN and overcomes the vanishing gradient issue. The existing neural networks have neurons in turn LSTM-RNN has blocks of memory connected to successive layers and the block contains the state of the block and output.

3.1 Dataset description

A standard IBM employee attrition dataset is considered in this work which has 35 features in it. These features are represented as 35 columns, out of which one column is “attrition”. From the remaining 34 features, “Standard hours” and “Employee count” are two features that are constant for all employees. So these features are not considered in the attrition process. The remaining 32 features in the dataset are listed below:

Age, business travel, daily rate, department, distance from home, education, education field, employee number, environment satisfaction, gender, hourly rate, job involvement, job level, job satisfaction, marital status, monthly income, monthly rate, number of companies worked, over 18, overtime, percent salary hike, performance rating, relationship satisfaction, stock option level, total working years, training times last year, work-life balance, years at company, years in the current role, years since last promotion, years with the current manager.

3.2 Data preprocessing

Pre-processing is an important task in the prediction process as, in this stage, irrelevant and unwanted data is removed from the dataset. The advantage of removing unwanted data is, it reduces the processing time and improves the prediction accuracy. The performance evaluation becomes easy due to the negligence of processing null information. Irrelevant employee data that doesn't come under the feature set is removed in the pre-processing stage. The processing of null information may lead to the wrong prediction of employee attrition [21]. Two pre-processed functions namely, normalization and numeralization are included in this work. This pre-processing function is mathematically represented as shown in Eq. (1).

$$P_{re} = Q_p[Ch_n] \tag{1}$$

Where, P_{re} is a pre-processing output function, Ch_n represents clustered input data and Q_p is the pre-processing function which signifies as Eq. (2).

$$Q_p=[Q_{Num}, Q_{Nom}] \quad (2)$$

Q_{Num} is the numeralization and Q_{Nom} is the normalization.

3.2.1. Numeralization

The process of converting clustered string values and characters into numerical values in the pre-processed data is called numeralization. Mathematically it is represented as shown in Eq. (3).

$$\bar{\lambda}^{Nu} = Q_{Num}[Ch_n] \quad (3)$$

Where, $\bar{\lambda}^{Nu}$ represents the outcome of the numeralization function.

3.2.2. Normalization

Allocating data values between the range 0 to 1 and -1 to 1 and utilizing overall feature values is called Normalization. The normalization of data compresses the wide range to small range values using a log scale. It is expressed in Eq. (4).

$$\bar{\lambda}^{Nor} = \log(Ch_n) \quad (4)$$

Ch_n is the original value and $\bar{\lambda}^{Nor}$ signifies the normalized value.

3.3 Feature selection

After eliminating unwanted features from the dataset in the pre-processing stage, the model is trained on the feature subset and then validated further. The LSTM-RNN method needs to be trained for 2^n times with n features which is a process that takes time. The feature selection method needs to be aligned with the feature characteristics. In this work, a butterfly optimization algorithm based on Brownian motion is used for feature selection.

3.3.1. Brownian motion-based butterfly optimization algorithm

The butterfly optimization algorithm (BOA) is a meta-heuristic algorithm in which the behavioral patterns of butterflies that scavenge for food and companions are impersonated. A chemoreceptor, which are sensory receptor dispersed in the butterfly body to detect or smell the fragrance of flowers and food. The functioning of these receptors also includes locating its companion. The butterflies emit the fragrance with a certain intensity while moving between locations. This fragrance directs the movement of the butterflies in the BOA algorithm. Based on the fragrance intensity, if the butterfly finds any companion within the search space, it will move in that direction which is known as global search (GS) or exploration. If the fragrance of any butterfly is not detected, then the butterfly engages in exploitation by moving in an arbitrary position which is also known as local search (LS). The characteristics of butterflies are: Every butterfly is meant to emit a fragrance that draws another butterfly to them. Each butterfly will move in an irregular direction of the butterfly emitting the most fragrance. The objective functions' landscape influences or determines the stimulus intensity of a butterfly.

Each butterfly will have a unique fragrance and personality and that is stated based on the intensity of the stimulus. It is very likely possible to determine using the following Eq. (5).

$$Fr = S.\beta_{si}^{a} \quad (5)$$

Where Fr is the variation in fragrance in each iteration, S indicates the sensor modality, the si represents the stimulus intensity and a represents the exponent of power that is reliant on the sensory modality. The range of a and S are [0, 1] and β is the Fitness function in Eq. (5). At every iteration of the butterfly, the butterfly moves in every possible direction in search space which changes its position. After that, each butterfly's fitness value (β) is determined and modified.

The two movements GS and LS phases of BOA are considered where the butterfly moves towards other butterfly in the GS phase in search of the best solution and it is mathematically represented as shown in Eq. (6). The LS phase is represented as shown in Eq. (7).

$$Y_i^{T+1} = Y_i^T + (R^2G^* - Y_i^T)Fr_i \quad (6)$$

where, Y_i^T represents vector solution for i_{th} butterfly iteration T, R represents an arbitrary number in the range [0, 1]. G^* significates the best solution for the current iteration, Fr_i represents fragrance of i_{th} butterfly.

$$Y_v^{T+1} = Y_i^T + (R^2Y_j^T - Y_v^T)Fr_i \quad (7)$$

where, Y_j^T and Y_v^T represent the j_{th} and v_{th} solution room of the butterfly, R represents the random number in the range[0, 1]. This random replacement can lead to limitations like the replacement of the undesirable individual at random and a lack of exploitation. The slow convergence issue might result from these limits. To balance this Brownian motion is combined with the optimization and BR_M is represented as follows in Eqs. (8), (9), (10), and (11).

$$BR_M = \aleph \times R_d(.) \times \emptyset_I \tag{8}$$

Where,

$$\aleph = \sqrt{\frac{\gamma}{\mu}} \tag{9}$$

$$\mu = 100 \times \gamma \tag{10}$$

$$\lambda_L = \frac{1}{\aleph\sqrt{2\pi}} \exp\,(-\frac{(dimension-agent)^2}{2\aleph^2}) \tag{11}$$

Where, γ represents the motion time period (seconds), μ represents the no. of sudden motions. R_d is the random probability constant and λ_L is the movement operator with a value of 0.5. The improved version of GS and LS phase with the involvement of BM for the optimization is represented in Eqs. (12) and (13).

$$Y_i^{T+1} = Y_i^T + (BR_M^2 G^* - Y_i^T)Fr_i \tag{12}$$

$$Y_v^{T+1} = Y_i^T + (BR_M^2 Y_j^T - Y_v^T)Fr_i \tag{13}$$

Once the termination norm is met, the butterfly moment comes to an end. The termination is norm is determined based on the number of served iterations. Based on the fitness value, the algorithm gives better result. The algorithm produces the best result based on the fitness value. Hence the selected features are represented mathematically in Eq. (14).

$$S(F_e) = \{\alpha_1, \alpha_2, \alpha_3 \dots . \alpha_t\} \tag{14}$$

Where, α is a feature in the feature set $S(F_e)$.

3.4 Classification

A long short-term memory (LSTM) based recurrent neural network is considered for classification purposes to which the selected features are given as input. The LSTM is incorporated in RNN to represent the prediction outcomes in time series and to attain better classification results. LSTM is combined with RNN which can possess long short memory as well as long term dependency with the backpropagation nature of RNN and overcomes the vanishing gradient problem. The LSTM-RNN has memory blocks that are connected to the successive layers and the block contains its block's state along with the output.

3.4.1. Long short term memory (LSTM) based recurrent neural network (RNN):

RNN is a neural network type in which the output from the previous state is taken as input for the current state. The important characteristic of RNN is, that it consists of hidden layers (HL), where these layers maintain a sequence of data and a set of weights and biases in the network. By iterating the successive sequence from $t = 1$ to T, the LSTM-RNN is implemented on the input data $I_p = \{\alpha_1, \alpha_2, \alpha_3 \dots . \alpha_t\}$ which includes a vector sequence that is hidden $\hbar_{hid} = \{\hbar_1, \hbar_2, \hbar_3, , \dots, \hbar_t\}$ and the output vector sequence $O_{out} = \{O_1, O_2, O_3, \dots, O_n\}$. Hence the HL can be represented as shown in Eqs. (15), (16), (17), and (18).

$$\hbar_{hid} = \partial_{act}[W_{\alpha\hbar}\alpha_t + W_{\hbar\hbar}\hbar_{t-1} + B_a] \tag{15}$$

Where, $W_{\alpha\hbar}$ is input hidden weight matrix, B_a is a bias vector, ∂_{act} represents hidden Activation function (AF).

$$F(X) = X \times Sigmoid(\lambda X) \tag{16}$$

Where, λ is a trainable parameter in the concerned training model.

This function activates the output layer OL and is represented as

$$O_t = \sigma_S[W_{\alpha o}\hbar_t + B_a] \tag{17}$$

$$\alpha = W_{\alpha o}\hbar_t + B_a \tag{18}$$

Eq. (15) is used to determine the function of sigmoid in the following Eq. (19).

$$\sigma_S(\alpha) = \frac{1}{1+\varepsilon^{-\alpha}} \tag{19}$$

Using the difference between the value of actual (α) and predicted values ($\hat{\alpha}$), the error loss is calculated as shown in Eq. (20).

$$Er = (\alpha - \hat{\alpha})^2 \tag{20}$$

The proposed approach predicts the exact value if the error value is 0 i.e., (if $Er = 0$). If the error value is not equal to zero, then there may be chances of backpropagation while updating the weight values. Therefore, employee attrition is predicted correctly without any prediction errors.

Table 1. Parameter settings for proposed BM-BOA-based LSTM-RNN

Parameter	Value
Method	BM-BOA
λ_L	0.5
S	[0,1]
R	[0,1]
LSTM-RNN	
Epochs	100
Activation function	tan H
Learning rate	0.01
Optimizer	Adam
Number of layers	10
Batch size	32

3.5 Attrition process

The network examines the employee data if that specific employee comes under the attrition process. Making the network usage a point of focus, similar threshold values are determined. If the threshold value is higher, the attrition process needs to be carried out which determines that the network of an employee is larger. The parameter settings for predicting the employee attrition rate for the BM-BOA and the LSTM-RNN network are given in Table 1. Employee attrition is the phenomenon of maintaining former employees and current employees on the same network by making some appealing offers and preventing them to change for another network. The attrition process of an employee is ignored when the network usage is below the threshold value.

4. Results

In the proposed research article, the BM-BOA based LSTM-RNN model for the prediction of employee attrition rate is evaluated on Python 3.7 environments with configuration of 8GB RAM, i9 Intel core processor and the operating system of Windows 10. The performance of the model is analyzed using accuracy, f-score, precision and recall on the IBM HR dataset. The feature selection of 32 features has been obtained in this model. Accuracy is the most important performance measure that is utilized in the LSTM, where it is the ratio of correctly predicted observations from the total observations. Precision is the similarity measure of the obtained values and recall is the prediction of positive outcomes from the overall positive instances. Accuracy, precision and recall are mathematically defined in the Eqs. (21), (22), and (23).

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \tag{21}$$

$$\text{Precision}=\frac{TP}{TP+FP} \tag{22}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{23}$$

F-score is determined as the harmonic mean of the model's recall and precision, and its mathematical equation is followed in Eq. (24).

$$\text{F-score}=\frac{2TP}{2TP+FP+FN} \tag{24}$$

4.1 Quantitative analysis

The efficacy of the proposed LSTM-RNN model is evaluated on IBM HR dataset which consists of 35 features of corporate employees. Various employee attrition techniques were used in the existing research, out of which two prediction techniques were compared to the proposed LSTM-RNN in terms of optimization in this section. Table 2 represents the effectiveness of the existing weighted forest optimization algorithm (WFO) [22] and particle swarm optimization (PSO) [23] compared with the proposed brownian motion-based butterfly optimization algorithm (BM-BOA).

The results are being evaluated for the feature selection optimizers with the combination of RNN, LSTM and LSTM-RNN as shown in Figs. 2, 3, and 4 for the verification of the proposed combination of

Table 2. Simulation results of existing optimization techniques with proposed BM-BOA.

Classifiers	Optimizers	Accuracy (%)	Recall (%)	F-score (%)	Precision (%)
RNN	WFO	88.56	64.96	67.41	69.98
	PSO	91.20	68.31	70.63	73.33
	BM-BOA	94.99	73.00	77.75	79.85
LSTM	WFO	88.99	65.39	68.09	70.03
	PSO	91.72	68.68	71.53	73.96
	BM-BOA	95.28	73.91	78.58	79.94
RNN + LSTM	WFO	89.25	66.35	68.56	70.52
	PSO	92.45	68.88	71.87	74.54
	BM-BOA	96.68	96.62	96.62	96.64

Received: August 6, 2023. Revised: October 11, 2023.

Table 3. Comparative analysis of various classifiers on IBM HR dataset based on performance metrics

Methods	IBM HR Dataset Details	Accuracy (%)	Recall (%)	F-score (%)	Precision (%)
Logistic regression [12]	Features – 35 Workers – 1470 Type – Numeric and Categorical Scale – Ratio and Nominal	81	N/A	N/A	N/A
NELT [18]		94.52	94.5	94.5	94.5
DL [19]		94.52	94.52	94.52	94.58
IG-RF [20]		89.2	88.6	88.2	87.8
Proposed LSTM-RNN		**96.68**	**96.62**	**96.62**	**96.64**

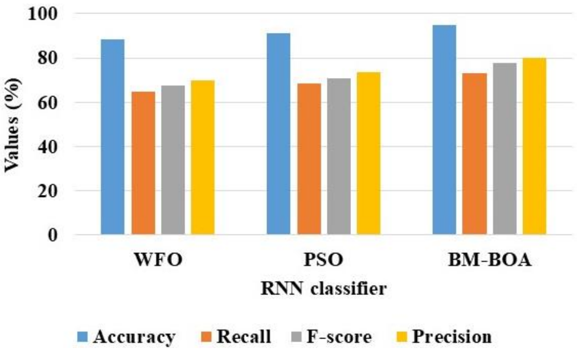

Figure. 2 Performance analysis of RNN on different optimizers

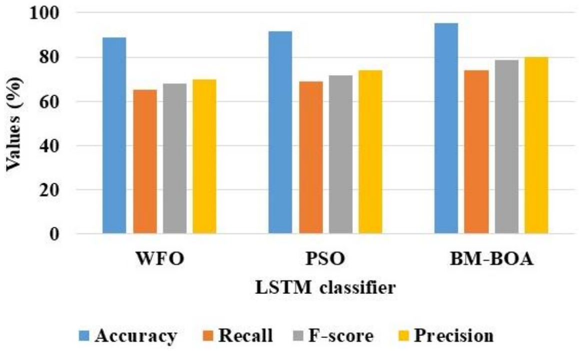

Figure. 3 Performance analysis of LSTM on different optimizers

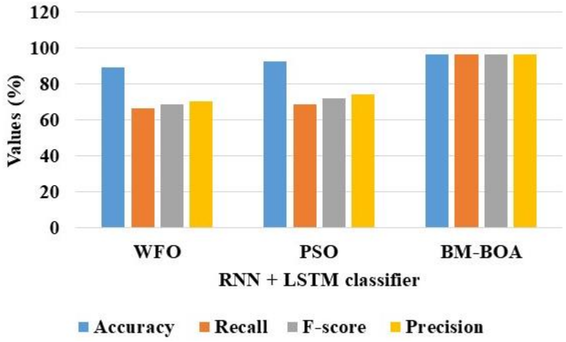

Figure. 4 Performance analysis of LSTM+RNN on different optimizers

LSTM-RNN to be better than the existing classifiers of LSTM and RNN.

The WFO algorithm was too slow in predicting once the algorithm was trained completely. Where the proposed optimization prediction accuracy is high.

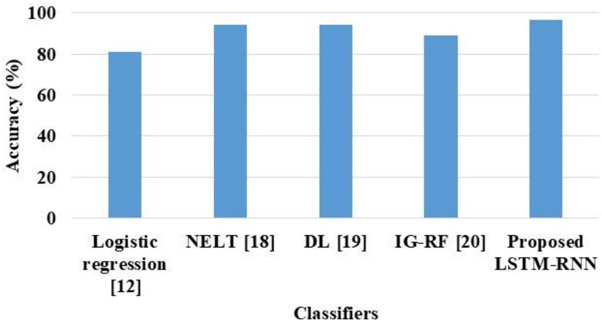

Figure. 5 Performance analysis of accuracy on different classifiers

The iteration is low in PSO which has been overcome with the proposed BM-BOA.

5. Comparative analysis

The comparative analysis between the proposed recurrent neural network-based LSTM to the existing logistic regression [12], NELT [18], DL [19] and IG-RF [20] classifier model in terms of accuracy, recall, F-score, and precision is shown in Table 3. The existing Gaussian Naïve Bayes approach failed to consider employees' intentions over attrition. This limitation has been overcome in this proposed method by classifying employee features into different groups. The logistic regression classifier failed to predict employee attrition with relevant features which resulted in more processing time.

This limitation has been overcome in this proposed approach by using the BM-BOA for the selection of relevant attributes for employee attrition prediction. The proposed model is validated using the IBM HR dataset. The accuracy was improved to 96.68%, recall of 96.62%, F-score of 96.62%, and precision of 96.64%. The graphical representation of accuracy is shown in Fig. 5.

6. Conclusion

In this paper, a novel classification technique for classifying employee attrition rates has been developed on the IBM HR dataset containing employee data with various features. The method uses an LSTM classifier to classify employee attrition and group them into attrition employees or non-attritional employees. The proposed algorithm on the other hand combines the traditional machine learning algorithms and RNN. The proposed model outperformed the existing classification techniques on employee attrition with suitable optimization with the LSTM classifier. A brownian motion-based butterfly optimization algorithm is used for feature selection in this work. The performance metrics show an accuracy of 96.68%, recall of 96.62%, F-score of 96.62%, and precision of 96.64%. The proposed approach remained the most reliable and robust model. In the future, the prediction of employee attrition can be done by considering various behavior patterns of an employee like adaptability, dependability, teamwork, etc.

Notations

Notation	Description
P_{re}	Pre-processing output function
Ch_n	Clustered input data
Q_p	Pre-processing function
Q_{Num}	Numeralization
Q_{Nom}	Normalization
$\bar{\lambda}^{Nu}$	Outcome of the numeralization function
$\bar{\lambda}^{Nor}$	Normalized value
Fr	Variation in fragrance in each iteration
S	Sensor modality
si	Stimulus intensity
a	Exponent of power
Y_i^T	Vector solution for i_{th} butterfly iteration T
R	Arbitrary number in the range [0, 1]
G^*	Best solution for the current iteration
Fr_i	Fragrance of i_{th} butterfly
γ	Motion time period
μ	Number of sudden motions
R_d	Random probability constant
λ_L	Movement operator
α	Feature in the feature set $S(F_e)$
$W_{\alpha\hbar}$	Input hidden weight matrix
B_a	Bias vector
∂_{act}	Hidden Activation function
TP	True positive
FP	False positive
TN	True negative
FN	False negative

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

[1] N. E. Rayes, M. Fang, M. Smith, and S. M. Taylor, "Predicting employee attrition using tree-based models", *International Journal of Organizational Analysis*, Vol. 28, No. 6, pp. 1273-1291, 2020.

[2] A. Qutub, A. A. Mehmadi, M. A. Hssan, R. Aljohani, and H. S. Alghamdi, "Prediction of Employee Attrition Using Machine Learning and Ensemble Methods", *International Journal of Machine Learning and Computing*, Vol. 11, No. 1, pp. 110-114, 2021.

[3] F. Ozdemir, M. Coskun, C. Gezer, and V. C. Gungor, "Assessing employee attrition using classifications algorithms", In: *Proc of the 2020 the 4th International Conference on Information System and Data Mining*, Hawaii HI, USA, pp. 118-122, May 2020.

[4] N. B. Yahia, J. Hlel, and R. C. Palacios, "From Big Data to Deep Data to Support People Analytics for Employee Attrition Prediction", *IEEE Access*, Vol. 9, pp. 60447-60458, 2021.

[5] I. Setiawan, S. Suprihanto, A. C. Nugraha, and J. Hutahaean, "HR analytics: Employee attrition analysis using logistic regression", *IOP Conference Series: Materials Science and Engineering*, Vol. 830, No. 3, p. 032001, 2020.

[6] S. K. M. Tharani and S. N. V. Raj, "Predicting employee turnover intention in IT&ITeS industry using machine learning algorithms", In: *Proc. of 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, pp. 508-513, November 2020.

[7] S. Dutta and S. K. Bandyopadhyay, "Employee attrition prediction using neural network cross validation method", *International Journal of Commerce and Management Research*, Vol. 6, No. 3, pp. 80-85, 2020.

[8] R. Joseph, S. Udupa, S. Jangale, K. Kotkar, and P. Pawar, "Employee Attrition Using Machine Learning And Depression Analysis", In: *Proc. of 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, pp. 1000-1005, May 2021.

[9] R. Eduvie, J. C. Nwaukwa, F. Uloko, and E. Taniform, "Predicting Employee Attrition using Decision Tree Algorithm", *Global Scientific Journals*, Vol. 9, No. 9, pp. 1305-1318, 2021.

[10] M. Gupta, N. Mandowara, K. Namdeo, and R. Patel, "Employee Attrition Rate Prediction", *International Research Journal of Modernization in Engineering Technology and Science*, Vol. 02, No. 04, pp. 1009-1013, 2020.

[11] F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. D. Luca, "Predicting Employee Attrition Using Machine Learning Techniques", *Computers*, Vol. 9, No. 4, p. 86, 2020.

[12] S. N. Zangeneh, N. S. Gharneh, A. A. Nezhad, and S. H. Zolfani, "An Improved Machine Learning-Based Employees Attrition Prediction Framework with Emphasis on Feature Selection", *Mathematics*, Vol. 9, No. 11, p. 1226, 2021.

[13] N. Jain, A. Tomar, and P. K. Jana, "A novel scheme for employee churn problem using multi-attribute decision making approach and machine learning", *Journal of Intelligent Information Systems*, Vol. 56, No. 2, pp. 279-302, 2021.

[14] A. B. W. Ali, "Prediction of Employee Turn Over Using Random Forest Classifier with Intensive Optimized Pca Algorithm", *Wireless Personal Communications*, Vol. 119, No. 4, pp. 3365-3382, 2021.

[15] X. Cai, J. Shang, Z. Jin, F. Liu, B. Qiang, W. Xie, and L. Zhao, "DBGE: Employee Turnover Prediction Based on Dynamic Bipartite Graph Embedding", *IEEE Access*, Vol. 8, pp. 10390-10402, 2020.

[16] S. Kakad, R. Kadam, P. Deshpande, S. Karde, and R. Lalwani, "Employee attrition prediction system", *International Journal of Innovative Science, Engineering & Technology*, Vol. 7, No. 9, pp. 57-63, 2020.

[17] M. Pratt, M. Boudhane, and S. Cakula, "Employee attrition estimation using random forest algorithm", *Baltic Journal of Modern Computing*, Vol. 9, No. 1, pp. 49-66, 2021.

[18] M. S. Alshiddy and B. N. Aljaber, "Employee Attrition Prediction using Nested Ensemble Learning Techniques", *International Journal of Advanced Computer Science and Applications*, Vol. 14, No. 7, pp. 932-938, 2023.

[19] S. M. Arqawi, M. A. A. Rumman, and E. A. Zitawi, "Predicting Employee Attrition and Performance Using Deep Learning", *Journal of Theoretical and Applied Information Technology*, Vol. 100, No. 21, pp. 6526-6536, 2022.

[20] S. F. Sari and K. M. Lhaksmana, "Employee Attrition Prediction Using Feature Selection with Information Gain and Random Forest Classification", *Journal of Computer System and Informatics (JoSYC)*, Vol. 3, No. 4, pp.410-419, 2022.

[21] R. Sudharsan and E. N. Ganesh, "A Swish RNN based customer churn prediction for the telecom industry with a novel feature selection strategy", *Connection Science*, Vol. 34, No. 1, pp. 1855-1876, 2022.

[22] S. Porkodi, S. Srihari, and N. Vijayakumar, "Talent management by predicting employee attrition using enhanced weighted forest optimization algorithm with improved random forest classifier", *International Journal of Advanced Technology and Engineering Exploration*, Vol. 9, No. 90, pp. 563-582, 2022.

[23] F. Kılıç, Y. Kaya, and S. Yildirim, "A novel multi population based particle swarm optimization for feature selection", *Knowledge-Based Systems*, Vol. 219, p. 106894, 2021.