International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Employee Attrition Rate Prediction Using Improved Sparrow Search Algorithm-based Deep Belief Network

Subrahmanya H M*[1], Dr. Shiva Prakash T[2]

Submitted: 27/01/2024 **Revised**: 05/03/2024 **Accepted**: 13/03/2024

Abstract: Employee attrition is a natural procedure by which employees leave the workspace because of uncertain factors like retirement, resignation for personal reasons, and are not replaced immediately. Decision-making plays a significant role in management and it is the most essential element in the planning process. However, Employee attrition is regarded as a well-known issue that requires the right decisions from administration to provide highly qualified employees. Therefore, the Improved Sparrow Search Algorithm-based Deep Belief Network (ISSA-DBN) is proposed to predict whether the employee is leaving or staying at the company using Deep Learning (DL).SSA is improved by using chaos mapping to generate increased population diversity, Adaptive inertia weight for updating the finder position, and an Adaptive t-distribution approach to solving the premature convergence. Initially, the IBM HR dataset is utilized to evaluate the proposed approach, and min-max normalization is established to improve the model performances. The Adaptive Synthetic Sampling technique (ADASYN) is used to balance the imbalanced data. Then, the ISSA is performed to select the appropriate features. Finally, DBN is employed for employee attrition rate prediction. The proposed ISSA-DBN achieves a high accuracy of 0.99 compared to existing techniques like Deep Neural Networks (DNN), max out Logistic Regression (LR), and deep data-driven technique using a Voting Classifier (VC) respectively.

Keywords: *Adaptive Synthetic Sampling technique, Chaos mapping, Deep Learning, Employee Attrition Rate, Improved Sparrow Search Algorithm*

1. Introduction

Employee attrition is a drop in the amount of workers of an organization where the employees leave the business retired or voluntarily. The greatly efficient employees are regarded as the most significant asset in an organization. Accommodating the high-performance employees or the most marketable is a great challenge in various organizations. The issue of employee attrition has gained popularity in various organizations due to negative effects on numerous subjects ranging from the performance of the organization and efficiency to progress the disturbance in projects and long-term growth approach [1]. Attrition refers to when an employee left the job for personal reasons and moves out immediately from one to another organization. Employee turnover has two kinds: first is involuntary where employees are fired or retired by the companies, and voluntary is the second, where employees left their jobs because of personal reasons [2]. In an organization, Human Resources (HR) plays a vital role in each aspect of HR function containing training, recruiting, retention, development, and compensation. In the business world, employee attrition has great attention in HR context analytics [3]. Moreover, when certain employees leave an organization, its project continuity, productivity, and growth strategies organizational turnover are affected [4] [5]. The experienced people's early retirement is 13% according to the Ministry of Trade Industry and Energy, whereas the new employee's early retirement is 66% respectively [6]. By calculating the rate of attrition, the factors and causes are identified that are required to be addressed for eliminating attrition. The rate of attrition is measured by splitting the amount of employees who have left the company by the average employee number over a certain time [7]. Simultaneously, high employee turnover increases the costs of advertising, training, hiring, and firing [8]. The turnover of voluntary is further split into functional and dysfunctional which defines high and low-performing workers [9]. Leadership plays a significant role in nurturing and shaping the desired behavior of employees [10] [11]. Without the performance of individuality, there is no economic sector, organizational, or team performance [12]. Approaches like predictive analysis accomplish the features depending on future research data that is employed to predict employee performance and attrition by utilizing different Machine Learning (ML) and DL approaches [13]. It is necessary to build a robust and accurate approach that predicts the attrition risk depending on quantitative factors which enables the HR and company leadership to generate proactive action for plan or retention for succession [14]. Data mining techniques contain three procedures which are pre-processing, feature selection, and classification [15]. However, Employee attrition is

[1] *Research Scholar, Department of Computer Science and Engineering, VVIT, Bengaluru, Affiliated to VTU, Belagavi, Karnataka, India.*
[2] *Professor,Department of Computer Science and Engineering, Vijaya Vittala Institute of Technology, Bengaluru, Affiliated to Visvesvarya Technological University, Belagavi, Karnataka, India Email: subrahmanya.hm@gmail.com ,shivprakasht@gmail.com*

regarded as a well-known issue that requires the right decisions from administration to provide highly qualified employees.

The primary contribution of the research is described as follows:

- ADASYN is used to balance the imbalanced data by minimizing bias established by imbalance class and adaptively shifts the decision boundary of prediction which maximizes learning for data distribution.
- ISSA is employed to select the most appropriate features. SSA is improved by using chaos mapping to generate increased population diversity, Adaptive inertia weight for updating the finder position, and an Adaptive t-distribution approach to solving the premature convergence.
- DBN is established to predict whether the employee is leaving or staying in the company and effectively manages non-linear relationships for modeling the complex factors in employee prediction.

The remaining part of the paper is organized as follows: Section 2 indicates a literature survey. Section 3 represents a proposed methodology. Section 4 illustrates a result of proposed methodology. Section 5 discusses a conclusion.

2. Literature Survey

The related work about employee attrition rate prediction is described along with its benefits and limitations.

Salah Al-Darraji*et al.* [16] implemented a DNN to increase the employee attrition prediction. The features of a dataset were evaluated to acknowledge their correlation with each other and to determine the significant features. The pre-processing was performed containing rescaling, data cleaning, and categorical data encoding to increase the performance of the model significantly. Then, the DNN was utilized to predict employee rate attrition. The cross-validation was performed in a DNN approach which provides a realistic performance and prevents overfitting issues. However, DNN needs a huge amount of labeled data due to their intricate architecture which demands extensive training samples for efficient learning.

Saeed Najafi-Zangeneh *et al.*[17] Presented a max-out LR to predict employee attrition. The m-max-out approach was employed for feature selection in the phase of pre-processing. Then, the LR technique's validity for attrition prediction was checked by evaluating the variations of parameters while they were trained across various bootstraps. This approach reduces dimensionality reduction and improves the model generalization by utilizing the m-max-out approach. However, non-linear issues were not addressed with LR due to it having the surface of linear decisions.

Nesrine Ben Yahia*et al*. [18] introduced a deep data-driven technique-based VC to detect and predict the intention of employees to leave or not. Initially, the scaling approach was employed to establish every feature on an identical scale by normalizing data that avoids outliers impacting the predictions. The Select Best and Recursive Feature Elimination (RFE) were performed to select the appropriate features. Then, the ensemble approach of DL and ML was employed for attrition prediction. The ensemble classifier was employed to integrate the classifiers and their predictions which increase robustness over a single classifier. However, this approach suffers from model interpretability due to complex and nonlinear input data transformations that hinder understanding the dynamic features' impact on prediction.

Ali Raza*et al*. [19] developed an Extra Tree Classifier (ETC) to predict employee attrition rate. The Employee Exploratory Data Analysis (EEDA) was employed to evaluate factors which cause employee attrition. The Synthetic Minority Oversampling Technique (SMOTE) was employed for enabling balanced data. Feature engineering approach was established to determine best-fit parameters by feature correlation for prediction. Finally, ETC was applied to predict employee attrition rate. This approach minimizes the complexity of model prediction due to an identical number of target distributions. However, ETC suffers from overfitting while managing small datasets and its nature was difficult to generate explicit insights into individual decision-making processes.

Ashish Kumar Biswas *et al*.[20]Ensemble learning to predict the intention to quit depending on selected features like Organizational Commitment (OC), Job Involvement (JI), Updating profiles on Job Portals (PJP), and Activities on Professional Networking Sites (APNS). The ensemble approaches like LR, Gradient Boosting (GB), and K-Nearest Neighbor (KNN) were employed to predict the attrition. This approach not only succeeded in achieving the attrition prediction but also determined the Stimulus-

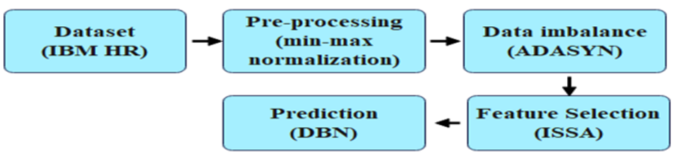

Fig 1. Block diagram for the proposed method

Organism-Response (S-O-R) theory for understanding the complex interaction among employee response, and stimuli on quitting intentions. However, the responders for this approach were only the employees who were professionals of Information Technology.

3. Proposed Methodology

The ISSA-DBN is proposed to predict employee attrition rate in this research. Initially, IBM HR dataset is employed to evaluate the proposed approach and min-max normalization is performed to improve the model performance. ADASYN is used for data imbalance issues and ISSA is established to select the appropriate features. Finally, DBN is used for employee attrition rate prediction. Figure 1 represents a block diagram for proposed method.

3.1. Dataset

The IBM HR [21] is used as a benchmark dataset to evaluate the proposed approach in this research. It has 35 columns for 1470 employees and one of the columns is attrition which is a target classifier output. Attrition indicates the decision of employees: yes (leave the company) or No (stay at company). The other34 columns are represented as features. The employee count and standard hours are the two features which are constant for every employee hence it is removed from the features. Other features are indicated as business travel, age, department, education, distance from home, education field, gender, employee number, and so on. These collected data are fed into the data pre-processing phase.

3.2. Pre-processing

After data acquisition, min-max normalization [22] is performed which is one of the commonly used normalizations. It implements the data's linear transformation and minimizes the feature value to a range among 0 and 1. It establishes this by subtracting the feature's minimum value from every data point and then dividing the outcomes by a feature value. The mathematical formula for this approach is represented in equation (1)

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where min and max indicates minimum and maximum values respectively. The preprocessed data are passed through data imbalance process.

3.3. Data Imbalance

Once the data is preprocessed, the data imbalance process is established using ADASYN [23] to balance the imbalanced data. Undersampling and Oversampling are the two standard re-sampling approaches. The samples of a majority class are minimized utilizing under sampling whereas the samples of a minority class are performed employing oversampling. Here, the ADASYN oversampling technique is employed to balance the data and generate more accurate and exact prediction findings. It generates synthetic data and its primary benefits are not copying the same minority data and providing more data for learning. The main aim of ADASYN is to employ a weighted distribution for various minority classes depending on their learning difficulties with large synthetic data established for minority classes. It increases a distribution of data in two ways: initially, it minimizes bias affected by imbalance class, and second, it pushes adaptively the decision boundary of classification toward challenging samples. It effectively balances each class sample and solves the imbalance issue in a dataset. ADASYN employs a Nearest-neighbor technique to include additional imitation instances for the minority class exhibited for balancing the dataset. After balancing data, the feature selection is established for selecting features.

3.4. Feature Selection

After performing the data imbalance, the ISSA selects the appropriate features for employee attrition rate prediction. ISSA iteratively determines subsets of attributes that select the most appropriate features and refines the search space for increasing prediction accuracy. This dynamic technique optimizes the model by adjusting adaptively the set of features and enhances its ability to predict employee attribute rates. SSA [24] [25] is a bionic intelligence technique depending on influence of two various sparrow behaviors in searching food procedure. It has more benefits than other optimization like Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), and so on in global optimization and operational accuracy. Sparrows are divided into two groups while searching for food: the finder and follower. Finder is the top-ranked fitness part in SSA. The finders are required to determine food and share the data regarding food with every sparrow. The rest of the sparrows were required to follow the finder's directions for food. When a certain sparrow is warned of danger, they transfer a message and move to safe places quickly for preventing predatory behavior. The distance correlation is employed as an objective function to measure the dependence among two random variables which captures non-linear relationships.

Consider there is N number of sparrows in the space of d-dimension, every position of sparrow and fitness value are $X = [x_1, x_2, \dots, x_n]$ and $F = f\ [x_1, x_2, \dots, x_n]$. The finders search food and direct the behavior of followers in SSA. The finder's location is updated which is expressed in equation (2)

$$X_{m,d}^{t+1} = \begin{cases} X_{m,d}^{t} \cdot \exp\left(-\frac{m}{\alpha . T}\right), & R_2 < ST \\ X_{m,d}^{t} + Q.L, & R_2 \geq ST \end{cases} \quad (2)$$

Where $X_{m,d}$ refers location data, $t+1$ indicates algebra of present evolutionary, T represents algebra of maximal evolutionary, both $\propto$ and Q determines uncertain figures, where Q refers distributed to N(0, 1), $\propto$ indicates among (0, 1], L indicates array with one row, D columns where each column value is 1, R_2determines alert value that is a numerical value among [0, 1], STillustrates secure value among [0.5, 1]. When $R_2 < ST$it determines there are no dangers at this neighbor's time, hence sparrows with greater fitness acquire a prosperous food. The position of follower's updated equation is expressed in equation (3)

$$X_{m,d}^{t+1} = \begin{cases} Q.\exp\left(\frac{X_{worst}^{t} - X_{m,d}^{t}}{m^2}\right), & m > \frac{n}{2} \\ X_{best}^{t+1} + \left|X_{m,d}^{t} - X_{best}^{t+1}\right| A^{+}.L, & m \le \frac{n}{2} \end{cases}$$

(3)

Where X_{best}^{t+1} and X_{worst}^{t}indicates contrary meaning that has worst and best position data at corresponding times of evolution. A represents matrix with one row and d columns, every components is ± 1, and $A^{+} = A^{T}(AA^{T})^{-1}$ is satisfied. The amount of early-warning sparrows is slightly greater than one-tenth population which is expressed in equation (4)

$$X_{m,d}^{t+1} = \begin{cases} X_{best}^{t} + \beta\left(X_{m,d}^{t} - X_{best}^{t}\right), & f_m \ne f_g \\ X_{m,d}^{t} + K.\left(\frac{\left|X_{m,d}^{t} - X_{worst}^{t}\right|}{(f_m - f_w) + e}\right), & f_m = f_g \end{cases}$$

(4)

Where β and K represents random numbers, β indicates random number to N (0, 1) that is a step size, K has [-1, 1], e refers infinitely small number which avoids dividend from having meaningless. While $f_m \ne f_g$, it represents certain individuals to leave the food range seeking and the danger is serious. Whether $f_m = f_g$ determines certain individuals among the mid position which has threat aware and this sparrow population prevent from risk and move to secure place.

3.4.1. Improved Sparrow Space Algorithm

Here, the SSA is improved by using chaos mapping approach to generate increased population diversity and high convergent accuracy. Adaptive inertia weight is established for updating finder's position. Adaptive t-Distribution approach is performed to solve the premature convergences which are described below.

3.4.2. Chaos Mapping Approach

Chaos is a primary phenomenon in nonlinear systems and is employed to intelligent swarm optimization technique because of pseudo-random and ergodic features. While SSA is initialized, the sparrow population is generated randomly, it will cause issue like low convergent accuracy and minimized population diversity. Hence, chaotic mapping is utilized to generate the population which provides better outcome. The most common chaotic mapping contains tent mapping, logistic mapping, Bernoulli mapping, and so on. The logistic mapping distribution points on both sides are greater and density is low in a middle which minimizes the speed of operations and other performance indicators. Moreover, tent mapping distribution is more uniform than of logistic mapping which has high convergence speed but has irregular periods and is prone to fixed points issues. The Bernoulli distribution mapping values are more uniform. It has no issues at fixed points and has faster convergence speed. Hence, Bernoulli mapping is utilized in SSA to manage with population defects which is indicated in equation (5)

$$Z_{k+1} = \begin{cases} \frac{Z_k}{1-\lambda}, & Z_k \in (0, 1-\lambda] \\ \frac{Z_k - 1 + \lambda}{\lambda}, & Z_k \in (1-\lambda, 1] \end{cases} \quad (5)$$

3.4.3. Weight Strategy

From equation (2) that while $R_2 \ge ST$, the finder moves towards set of distribution by technique and its value is converged to a optimal value. Whether $R_2 < ST$, the change of the position of finder is not effective which is expressed in equation (6)

$$y = \exp\left(\frac{-x}{\propto .T}\right) \quad (6)$$

Where x indicates present evolution number, T represents threshold set of upper evolutionary, and y refers finder's location of variation range. As x improves, y is minimized gradually, and its value becomes (0, 0.4). While x is 0, y represents uneven, and their probability of one id greater. As x enhances, y has uniform distribution hence, while $R_2 < ST$, the individual's variation range in the technique is minimized and becomes 0. This search approach increases the ability of SSA's local search. However, it was difficult to achieve an optimal value in following iterations. For further enhancing SSA's performance of optimization, an adaptive inertia weight is established for updating finder's new position which is expressed in equation (7)

$$X_{m,d}^{t+1} = \begin{cases} X_{m,d}^{t} \cdot \exp\left(-\frac{m}{w,\alpha,T}\right), & R_2 < ST \\ X_{m,d}^{t} + Q.L, & R_2 \geq ST \end{cases} \quad (7)$$

Where $w = 1 - \frac{\exp\left(\frac{t}{T}\right)-1}{\exp(1)}$ (8)

Where w indicates adaptive interia weight and its value is minimized gradually along with an increased t. While the t in the approach changes by including one, the w value is minimized gradually, strengthening the partial performance of SSA in the evolution phase like convergent accuracy.

3.4.4. Adaptive t-Distribution Approach

In the iteration of SSA, sparrows are moved nearer to the ideal individual step by step which causes a decrease in the sparrow numbers and the technique converges prematurely. Here, adaptive t-distribution is employed to optimize SSA for solving this problem. The t-distribution, Gaussian, and Cauchy distribution are conventional distributions. The t-distribution is a specific distribution which is transferred among Gaussian and Cauchy distribution, and it requires for modifying n parameter. While $n = 1$, t-distribution is similar to Cauchy C (0, 1), and t-distribution is similar to Gaussian N (0, 1) when n $\geq$ 30. The update equation for sparrow's location change is expressed in equation (9)

$$X_m^{t+1} = X_m^t + X_m^t . t\ (iter) \quad (9)$$

Where

$$t\ (iter) = \frac{Gam\left(\frac{iter+1}{2}\right)}{\sqrt{iter.\pi}\ Gam\left(\frac{iter}{2}\right)} \left(1 + \frac{x^2}{iter}\right)^{-\frac{iter+1}{2}}, -\infty < x < \infty$$

(10)

Where X_m^{t+1} refers new location data of m^{th} sparrow after a variation of t-distribution, X_m^t determines data location of m^{th} sparrow in i^{th} generation. $t\ (iter)$ indicates t-distribution and their freedom degree refers SSA's iteration number, and $Gam\ (x)$represents function of Gamma. A mutational update equation initially employs position of sparrow in present period. It includes adaptive t-distribution disturbance, and their degree of freedom modifies SSA iteration. The t-distribution is similar to Cauchy in early iteration stage, and the ability of global search is increased. In SSA's final iteration, a t-distribution becomes a Gaussian and capability of search technique is increased at this time that enables the accuracy of convergence to be more accurate. Thus, the above three techniques are established to improve the SSA hence that population diversity, and search algorithm is improved to optimize the performance for employee attrition. Then, the selected features are fed into prediction process.

3.5. Prediction

After selecting features, the DBN is utilized to predict the employee attrition rate whether the employee is leaving or staying at the company. The selected features are provided as input vectors and each feature correlates to a node in input layer of DBN. It employs various layers of hidden nodes to learn hierarchical features automatically from the input data which captures intricate pattern in features like performance and job satisfaction. DBN effectively manages non-linear relationships for modeling the complex factors in employee prediction. It contains series of individually trained Restricted Boltzmann Machines (RBMs) which are stacked on each other. The RBMs in a DBN is trained in a manner of unsupervised learning, where the process of training starts with an unsupervised phase. There are two processing layers in every RBM which are represented as hidden and visible layer in a DBN. The hidden layer captures hidden or latent representation whereas visible layer indicates observable features or entities in a data.

The DBN training includes iteratively training every RBM in a layer-wise manner. The initial RBM is trained utilizing visible layer as input and for next RBM, its activation of hidden layer become the visible layer. This Procedure will be continued till every RBM is trained. This pre-training layer wise assist to fix issues which is occurred while the network is set up initially with untrained, and connection weights of arbitrary. The generative stochastic neural networks are learned from the model of probability by utilizing unsupervised learning approaches. A huge amount of observable entities $(v_1, v_2, \ldots, v_i)$ enable the visible layer $(v)$of network, that is trained on the structure of unlabeled pattern that is fed into it, and a huge amount of unseen entities $(h_1, h_2, \ldots, h_j)$ and it has binary values, receive data from seen nodes which is able to build the patterns (h). All obvious nodes have a two-way symmetric weight matrix (S_{ij}) with bias (b_i) and (a_i)which is represented in equation (11)

$$R\ (v,h) = \sum_{nevis} \frac{(v_i - b_i)^2}{2\lambda_i^2} - \sum_{j\epsilon hid} a_j h_j - \sum_{ij} \frac{v_i}{\lambda_i} h_j S_{ij}$$

(11)

Where λ indicates Gaussian noise dispersion in i^{th} visible dimension. The learning procedure has difficulties if both the concealed and exposed units are Gaussians. To compute the quadratic coefficients "containment" terms that maintains the activity within reasonable bounds. The energy function is formulated in equation (12)

$$R\ (v,h) = \sum_{nevis} \frac{(v_i - b_i)^2}{2\lambda_i^2} + \sum_{j\epsilon hid} \frac{(h_j - a_j)^2}{2\lambda_i^2} - \sum_{ij} \frac{v_i h_j}{\lambda_i \lambda_j} S_{ij} \quad (12)$$

Training set data is employed to determine the hidden unit probability and indicates those predictions in equation (13)

$$M(h_j = 1) = l\ (a_j + \sum v_i w_{ij}) \quad (13)$$

The invisible variable $v'$is reconstructed at visible level with sample$h$. Then, ah' fresh set of hidden activations is collected which is expressed in equation (14)

$$M\ (v_i = 1) = l\ (b_i + \sum h_j s_{ij}) \quad (14)$$

The outcome of multiplying$v'$by h' from the outside is key to the (negative phase) solution. The law of Weight matrix is expressed in equation (15)

$$\Delta S_{ij} = \eta((v_i.h_j)_{data} - (v_i'.h_j')_{model}) \quad (15)$$

Where η represents learner speed and $(.)$ indicates logistic activation function. The changes are generated to b_i and $h_j$in equations (14) and (15) is represented in equations (16) and (17)

$$b = b + l\ (v - v') \quad (16)$$

$$a = a + l\ (h - h') \quad (17)$$

The logistic activation function is expressed in equation (18) that is utilized in every processing node.

$$\emptyset\ (x) = \frac{1}{1+e^{-x}} \quad (18)$$

This function generates an input value (x) and employs the transformation of logistic to squash output among 0 and 1. DBN effectively capture non-linear dependencies which provides accurate prediction on if an employee will stay or leave the company by encoding selected features into the input layer of network. This approach increases the processes of decision making in workspace management.

Table 1 indicates the Notation description

Table 1. Notation Description

Symbols	**Description**
$X_{m,d}$	location data
$t+1$	present evolutionary algebra
T	maximal evolutionary algebra
$\propto$ and Q	uncertain figures
L	array with one row
D	columns
R_2	alert value which is numerical value among [0, 1]
ST	secure value among [0.5, 1]
X_{best}^{t+1} and X_{worst}^{t}	contrary meaning that has worst and best position data at corresponding times of evolution
β and K	random numbers
e	infinitely small number which avoids dividend from having meaningless
$f_m \neq f_g$	certain individuals to leave the food range
$f_m = f_g$	certain individuals among the mid position
x	present evolution number
T	threshold set of upper evolutionary
y	finder's location of variation range
$R_2 < ST$	individual's variation range
w	adaptive inertia weight
X_m^{t+1}	new location data of m^{th} sparrow after a variation of t-distribution
X_m^t	data location of m^{th} sparrow in i^{th} generation
$t\ (iter)$	t-distribution their degrees of freedom refers SSA's iteration number
$Gam\ (x)$	Gamma function
(v)	visible layer network
(S_{ij})	two-way symmetric weight matrix
λ	Gaussian noise dispersion in i^{th} visible dimension
v'	invisible variable
h'	fresh set of hidden activations
η	learner speed
$(.)$	logistic activation function

4. Results

The ISSA-DBN performance is simulated by Python 3.8 environment, i9 Intel processor, 128GB RAM, and Windows 10 operating system. The performance metrics like accuracy, f1-score, recall, and precision are employed in this research. Accuracy refers to the amount of exact prediction divided by overall predictions. Precision determines exactly predicts the positive class. Recall evaluates model capability to determine all appropriate cases within dataset. F1-score indicates combination of both precision and recall. Their mathematical formulas are represented in equations (19) to (22)

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \times 100 \quad (19)$$

$$Precision = \frac{TP}{TP+FP} \quad (20)$$

$$Recall = \frac{TP}{TP+FN} \quad (21)$$

$$F1 - Score = \frac{2 \times TP}{2 \times TP+FP+FN} \times 100 \quad (22)$$

Where TP indicates True Positive, FPdetermines False Positive, TN represents True Positives, and FN illustrates False Negative.

4.1. Performance Analysis

The performance evaluation of ISSA-DBN is indicated in Tables 1 to 3. Table 1 represents different feature selection techniques using IBM HR dataset. The performance of PSO, GWO, Mayfly Optimization Algorithm (MOA), and SSA are compared with ISSA. Figure 1 represents the graphical representation of different feature selection technique. The obtained result indicates that ISSA achieves better accuracy of 0.99 compared to PSO, GWO, MOA, and SSA respectively.

Table 1. Different feature selection techniques

Performance metrics	PSO	GWO	MOA	SSA	ISSA
Accuracy	0.86	0.87	0.89	0.91	0.99
F1-score	0.85	0.86	0.87	0.90	0.98
Recall	0.84	0.85	0.88	0.89	0.95
Precision	0.83	0.85	0.89	0.87	0.97

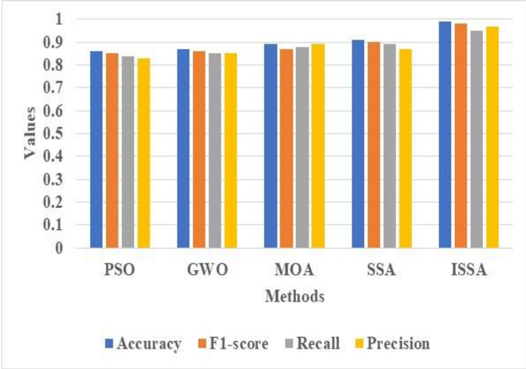

Fig 2. Graphical representation of different feature selection techniques

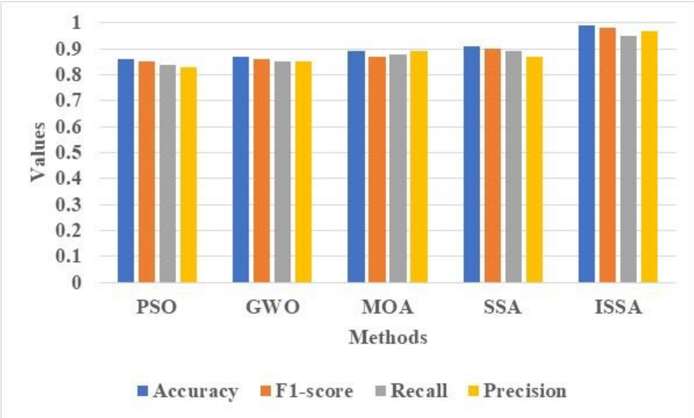

Table 2 indicates classification performances with default features. The performance of Gated Recurrent Unit (GRU), Multilayer Perception (MLP), Recurrent Neural Network (RNN), and Convolution Neural Network (CNN) are compared with DBN. Figure 3 illustrates graphical representation of classification performance with default features. The results show that DBN achieves better accuracy of 0.95 compared to existing methods.

Table 2. Classification performance with default features

Performance metrics	GRU	MLP	RNN	CNN	DBN
Accuracy	0.90	0.91	0.92	0.93	0.95
F1-score	0.89	0.89	0.90	0.91	0.94
Recall	0.87	0.90	0.91	0.90	0.92
Precision	0.88	0.88	0.89	0.91	0.93

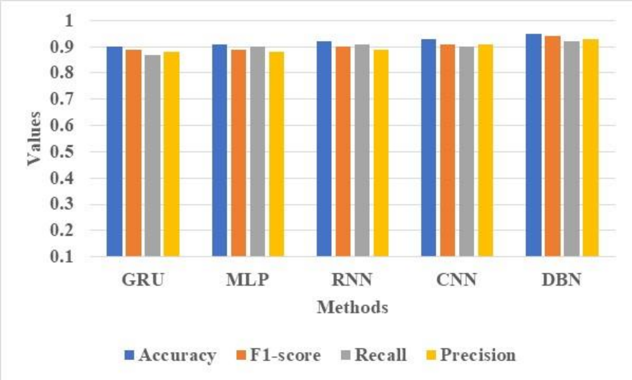

Fig 3. Classification performance with default features

Table 3. Classification performance with optimized features

Performance metrics	GRU	MLP	RNN	CNN	DBN
Accuracy	0.95	0.96	0.97	0.97	0.99
F1-score	0.94	0.95	0.96	0.95	0.98
Recall	0.93	0.94	0.96	0.96	0.95
Precision	0.94	0.95	0.95	0.94	0.97

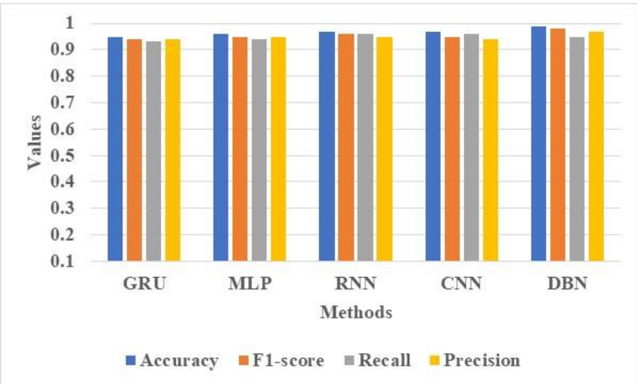

Fig 4. Graphical representation of classification performance with optimized features

4.2. Comparative Analysis

Table 4 represents comparative analysis with existing methods using IBM HR dataset. The existing techniques like DNN [16], Max-out [17], and VC [18] are compared with proposed ISSA-DBN. The proposed ISSA-DBN achieves greater accuracy of 0.99 whereas existing techniques have accuracy of 0.9116 in DNN, and 0.96 in VC respectively.

Table 4. Comparative analysis with existing techniques using IBM HR dataset

Performance metrics	DNN [16]	Max-out [17]	VC [18]	ISSA-DBN
Accuracy	0.9116	N/A	0.96	0.99
F1-score	0.91	0.56	N/A	0.98
Recall	0.91	0.82	N/A	0.95
Precision	0.9	0.43	N/A	0.97

4.3. Discussion

The advantage of ISSA-DBN and limitations of existing techniques are discussed in this section. The limitation of existing techniques like DNN [16] needs a huge amount of labeled data due to their intricate architecture which demands extensive training samples for efficient learning. Max-out VR [17] has non-linear issues were not addressed with LR due to it having the surface of linear decisions. VC [18] suffers from model interpretability due to complex and nonlinear input data transformations that hinder the understanding of the dynamic features' impact on prediction. The proposed approach overcomes these existing techniques' limitations. ISSA iteratively determines subsets of attributes that select the most appropriate features and refines the search space for increasing prediction accuracy. This dynamic technique optimizes the model by adjusting adaptively the set of features and enhances its ability.DNN effectively manages non-linear relationships for modeling the complex factors in employee prediction. Therefore, by combining these approaches, ISSA-DBN achieves greater accuracy of 0.99 compared to DNN, max out, and VC techniques.

5. Conclusion

In this research, the ISSA-DBN is proposed to predict whether the employee is leaving or staying in the company. SSA is improved by using chaos mapping to generate increased population diversity, Adaptive inertia weight for updating the finder position, and an Adaptive t-distribution approach to solving the premature convergence. ISSA iteratively determines subsets of attributes that select the most appropriate features and refines the search space for increasing prediction accuracy. DBN effectively manages non-linear relationships for modeling the complex factors in employee prediction. ISSA-DBN achieves greater accuracy of 0.99 compared to DNN, max out, and VC techniques. In future, improved prediction approach will be used to further enhance the prediction accuracy.

References

[1] Alsheref, F.K., Fattoh, I.E. and M Ead, W., 2022. Automated Prediction of Employee Attrition Using Ensemble Model Based on Machine Learning Algorithms. *Computational Intelligence and Neuroscience*, *2022*.

[2] Wild Ali, A.B., 2021. Prediction of employee turn over using random forest classifier with intensive optimized pca algorithm. *Wireless Personal Communications*, *119*(4), pp.3365-3382.

[3] Saha, L., Tripathy, H.K., Gaber, T., El-Gohary, H.

and El-kenawy, E.S.M., 2023. Deep churn prediction method for telecommunication industry. *Sustainability*, *15*(5), p.4543.

[4] Jain, N., Tomar, A. and Jana, P.K., 2021. A novel scheme for employee churn problem using multi-attribute decision making approach and machine learning. *Journal of Intelligent Information Systems*, *56*, pp.279-302.

[5] Park, J., Feng, Y. and Jeong, S.P., 2024. Developing an advanced prediction model for new employee turnover intention utilizing machine learning techniques. *Scientific Reports*, *14*(1), p.1221.

[6] Pustokhina, I.V., Pustokhin, D.A., Aswathy, R.H., Jayasankar, T., Jeyalakshmi, C., Díaz, V.G. and Shankar, K., 2021. Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms. *Information Processing & Management*, *58*(6), p.102706.

[7] Wang, X. and Zhi, J., 2021. A machine learning-based analytical framework for employee turnover prediction. *Journal of Management Analytics*, *8*(3), pp.351-370.

[8] Pekel Ozmen, E. and Ozcan, T., 2022. A novel deep learning model based on convolutional neural networks for employee churn prediction. *Journal of Forecasting*, *41*(3), pp.539-550.

[9] Kakulapati, V. and Subhani, S., 2023. Predictive Analytics of Employee Attrition using K-Fold Methodologies. *IJ Mathematical Sciences and Computing*, *1*, pp.23-36.

[10] Chowdhury, S., Joel-Edgar, S., Dey, P.K., Bhattacharya, S. and Kharlamov, A., 2023. Embedding transparency in artificial intelligence machine learning models: managerial implications on predicting and explaining employee turnover. *The International Journal of Human Resource Management*, *34*(14), pp.2732-2764.

[11] Lazzari, M., Alvarez, J.M. and Ruggieri, S., 2022. Predicting and explaining employee turnover intention. *International Journal of Data Science and Analytics*, *14*(3), pp.279-292.

[12] Naz, K., Siddiqui, I.F., Koo, J., Khan, M.A. and Qureshi, N.M.F., 2022. Predictive modeling of employee churn analysis for IoT-enabled software industry. *Applied Sciences*, *12*(20), p.10495.

[13] López-Cabarcos, M.Á., Vázquez-Rodríguez, P. and QuinoA-Pineiro, L.M., 2022. An approach to employees' job performance through work environmental variables and leadership behaviours. *Journal of Business Research*, *140*, pp.361-369.

[14] Bagheri, A., Newman, A. and Eva, N., 2022. Entrepreneurial leadership of CEOs and employees' innovative behavior in high-technology new ventures. *Journal of Small Business Management*, *60*(4), pp.805-827.

[15] Patel, K., Sheth, K., Mehta, D., Tanwar, S., Florea, B.C., Taralunga, D.D., Altameem, A., Altameem, T. and Sharma, R., 2022. RanKer: An AI-based employee-performance classification scheme to rank and identify low performers. *Mathematics*, *10*(19), p.3714.

[16] Al-Darraji, S., Honi, D.G., Fallucchi, F., Abdulsada, A.I., Giuliano, R. and Abdulmalik, H.A., 2021. Employee attrition prediction using deep neural networks. *Computers*, *10*(11), p.141.

[17] Najafi-Zangeneh, S., Shams-Gharneh, N., Arjomandi-Nezhad, A. and Hashemkhani Zolfani, S., 2021. An improved machine learning-based employees attrition prediction framework with emphasis on feature selection. *Mathematics*, *9*(11), p.1226.

[18] Yahia, N.B., Hlel, J. and Colomo-Palacios, R., 2021. From big data to deep data to support people analytics for employee attrition prediction. *IEEE Access*, *9*, pp.60447-60458.

[19] Raza, A., Munir, K., Almutairi, M., Younas, F. and Fareed, M.M.S., 2022. Predicting employee attrition using machine learning approaches. *Applied Sciences*, *12*(13), p.6424.

[20] Biswas, A.K., Seethalakshmi, R., Mariappan, P. and Bhattacharjee, D., 2023. An ensemble learning model for predicting the intention to quit among employees using classification algorithms. *Decision Analytics Journal*, *9*, p.100335.

[21] Dataset Link: https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset

[22] Silivery, A.K., Kovvur, R.M.R., Solleti, R., Kumar, L.S. and Madhu, B., 2023. A model for multi-attack classification to improve intrusion detection performance using deep learning approaches. *Measurement: Sensors*, *30*, p.100924.