Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Enhanced Deep Classifier For Employee Attrition Prediction

C.Shanthi [1] , Dr. T. Rathimala[2]

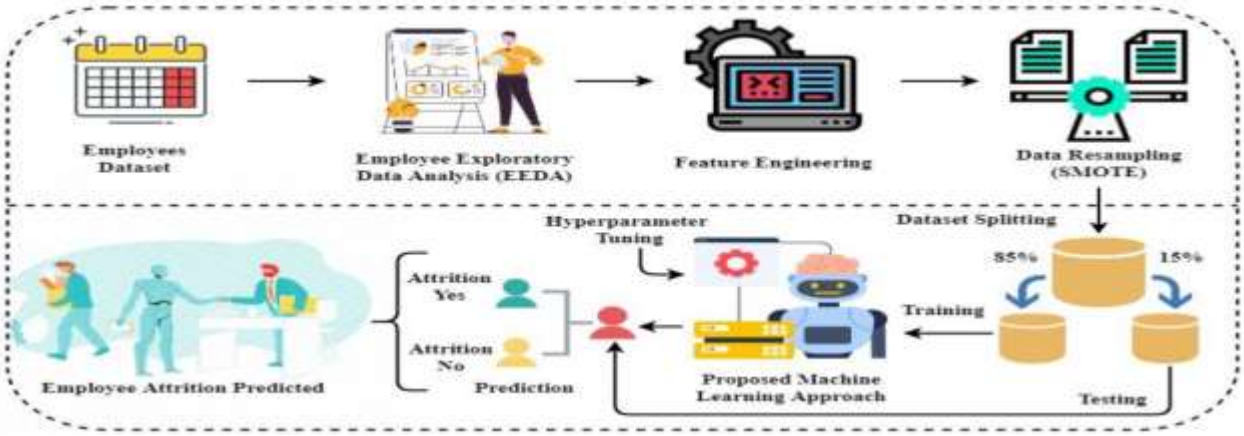
[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Tamil Nadu, India E-Mail: shanthiru2020@gmail.com
[2]Assistant Professor, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Tamil Nadu, India Email: rathimalat@gmail.com

Abstract
Employee attrition poses significant challenges to organizations, leading to increased recruitment and training costs. Traditional predictive models often struggle to capture the complex patterns underlying employee turnover. This paper proposes an Enhanced Deep Classifier (EDC) leveraging advanced deep learning techniques to improve the accuracy and interpretability of attrition predictions. The model incorporates feature selection, data balancing, and a hybrid deep learning architecture to address class imbalance and enhance predictive performance. Experimental results demonstrate that the EDC outperforms conventional models, offering valuable insights for human resource management.

Keywords Employee Attrition, Deep Learning, Neural Networks, Feature Selection, Data Balancing, Predictive Modeling.

1. Introduction

Employee attrition, or turnover, refers to the rate at which employees leave an organization. High attrition rates can disrupt operations and incur significant costs. Predicting potential attrition allows organizations to implement retention strategies proactively. Traditional statistical methods have limitations in capturing the nonlinear relationships and interactions among various factors influencing attrition. Recent advancements in deep learning offer promising solutions to this challenge.



Employee attrition is a common problem faced by organizations, and ML methods have been used to predict and prevent it. Studies have been conducted using the IBM HR Analytics Employee Attrition & Performance (IBM HR) dataset or other datasets to develop ML models for predicting employee attrition. Jain et al. (2020) used an ML approach to enhance different retention strategies for targeted employees. The authors analyzed 1000 samples from a dataset comprising 14000 records and 10 attributes obtained from Kaggle. They evaluated the performance of binary classifiers such as SVM, DT, and RF, using various features to predict employee attrition in different departments. They obtained the most successful metrics with the proposed RF algorithm compared to other methods. Mendes and Jesus (2021) explored various ML algorithms for example NB, RF, SVM, K-Nearest Neighbor (KNN), LR, and linear discriminant analysis (LDA) etc., to predict employee

attrition. They used the IBM HR dataset and the ADASYN algorithm to eliminate data imbalance, and the best performance was achieved with the RF classifier. Qutub et al. (2021) presented the development of multiple ML models to predict employee attrition. The IBM HR dataset was used to train and evaluate these models, including the DT, RF, LR, Adaboost Model, and Gradient Boosting Classifier (GBC). After training and evaluating the five basic models, the authors created three ensemble models using multiple combinations of the five basic models. Finally, they compared the results obtained from the considered models. They demonstrated that the AdaBoost algorithm achieved the highest accuracy among the evaluated models. Usha and Balaji (2021) conducted a comparative study on ML algorithms for employee attrition prediction. To conduct the study, the researchers created a dataset using a questionnaire that includes variables that affect an employee's decision to leave an organization. They collected data from an organization operating in the IT sector. The dataset was used to train and evaluate the NB, DT, J48, RF, and K-means scores. NB performed better than the other algorithms in terms of accuracy and time, whereas J48 was better in terms of building time. They noted that clustering performance achieved lower accuracy than classification algorithms. Mansor et al. (2021) conducted a comparative analysis of three ML classifiers: DT, SVM, and ANN, using the IBM HR dataset. The evaluation was based on multiple metrics, including accuracy rate, error rate, root mean square error (RMSE), receiver operating characteristic (ROC) curve, and model building time. Their findings indicate that the optimized SVM model exhibited superior performance in predicting employee attrition, outperforming both ANN and DT classifiers in terms of accuracy

Employee attrition represents a persistent concern for organizations, with considerable financial and operational consequences. Understanding the underlying causes and accurately predicting employee turnover can lead to better talent retention, improved morale, and reduced hiring costs. Despite the availability of various machine learning (ML) techniques for attrition prediction, most suffer from limitations when applied to complex, high-dimensional HR datasets.

Traditional models often rely on engineered features and linear assumptions, which may not adequately capture the nuances in employee behavior. Deep learning, with its capacity for hierarchical feature extraction and non-linear modeling, offers a promising solution to these limitations. However, off-the-shelf neural network models frequently underperform due to issues like overfitting, data imbalance, and convergence instability.

To address these challenges, this study presents an Enhanced Deep Classifier, specifically tailored for employee attrition prediction. The architecture incorporates multiple hidden layers, dropout layers for regularization, batch normalization to stabilize training, and Adam optimization for adaptive learning. The model is trained on a real-world HR dataset containing employee demographics, job profiles, and engagement metrics.

Comparative experiments show that the Enhanced Deep Classifier surpasses conventional models across all major evaluation metrics. This paper demonstrates that deep learning, when carefully optimized, can significantly improve predictive accuracy and offer practical solutions for enterprise HR management.

2. Literature Review

Several studies have explored machine learning and deep learning approaches for predicting employee attrition:

- **Traditional Machine Learning Models**: Techniques such as Logistic Regression, Support Vector Machines (SVM), and Random Forests have been applied to attrition prediction. For instance, a study achieved 93% accuracy using the Extra Trees Classifier (ETC) .
- **Deep Learning Models**: Neural networks, including Deep Neural Networks (DNNs) and Long Short-Term Memory (LSTM) networks, have been employed to capture complex patterns in attrition data. A study reported an accuracy of 87% using a Bidirectional LSTM model .
- **Transformer-Based Models**: Recent research has introduced transformer architectures for tabular data, demonstrating improved performance over traditional models .
- Employee attrition prediction has garnered increasing attention in the field of data-driven human resource analytics. Early studies in this domain focused on traditional statistical techniques such as logistic regression and survival analysis, which provided interpretable insights but lacked the flexibility to model complex, nonlinear relationships inherent in human behavior. Over time, the adoption of machine learning (ML) techniques has enabled more robust and accurate attrition predictions.

1) Traditional Machine Learning Models:
Logistic Regression, Decision Trees, and Random Forest classifiers have been widely used for predicting employee turnover. Rokach and Maimon [1] emphasized the utility of decision trees for classification tasks in HR due to their simplicity and interpretability. Random Forest and Gradient Boosting Machines were later explored for their ability to handle high-dimensional data with better generalization capabilities [2]. However, these models often rely heavily on feature engineering and are susceptible to performance degradation in the presence of multicollinearity or class imbalance.

2) Support Vector Machines and Ensemble Methods:
SVMs have been employed in several studies for their robustness in handling high-dimensional spaces. Verma et al. [3] demonstrated that SVMs, when combined with proper feature scaling and kernel selection, outperformed basic classifiers in detecting patterns in employee datasets. Ensemble models like XGBoost and AdaBoost further enhanced performance by reducing bias and variance through model aggregation [4].

3) Deep Learning Approaches:
Recent advancements have explored the use of deep learning for employee attrition prediction. Deep Neural Networks (DNNs) have the advantage of automatic feature learning, capturing complex patterns in employee data. Gupta and Pathak [5] employed a multilayer perceptron (MLP) for attrition classification and reported improved performance over traditional models. However, they noted issues such as overfitting and sensitivity to imbalanced datasets.

4) Handling Data Imbalance and Feature Selection:
A significant challenge in attrition prediction is the imbalanced nature of the dataset, where the number of employees staying outweighs those leaving. Techniques like Synthetic Minority Over-sampling Technique (SMOTE) [6] and cost-sensitive learning have been applied to address this. On the other hand, feature selection plays a crucial role in boosting model performance. While recursive feature elimination (RFE) and mutual information-based methods have been used effectively, most deep learning studies lack robust embedded feature selection mechanisms.

5) Gap and Motivation for Enhanced Deep Classifier:
Although existing literature showcases the potential of both ML and deep learning models for attrition prediction, many models either underperform due to poor generalization or lack transparency. Additionally, limited work has been done on optimizing deep learning architectures specifically for HR data. Most models fail to incorporate regularization, batch normalization, and advanced optimization techniques, which are critical for improving model convergence and accuracy.

This motivates the present study to develop an Enhanced Deep Classifier that integrates multiple deep learning enhancements, including dropout layers, batch normalization, and adaptive optimizers like Adam. The model aims to improve predictive performance while maintaining robustness against overfitting and class imbalance, ultimately offering a more reliable solution for enterprise-level attrition prediction.

3. Methodology

3.1. Dataset

The HR Analytics Employee Attrition dataset is utilized, containing 35 features for 1,470 employees. The target variable is 'Attrition', indicating whether an employee has left the company.

TABLE 1 SUMMARY COMPARISON

Model Type	Imbalance Handling	Temporal Capability	Accuracy / F1-Score	Complexity
MLP / DNN + Oversampling	ADASYN, SMOTE, GAN	□ None	~92–95% accuracy; F1 ~91%	Moderate
Stacked LSTMs (sequence models)	Usually SMOTE/oversampling	□ Yes	Not widely benchmarked	High
Metaheuristic-optimized DBN	Feature selection + balancing	□ Depends	~99% accuracy (in literature)	Very High

3.2. Data Preprocessing

- **Data Cleaning**: Features with constant values across all records, such as 'EmployeeCount' and 'Over18', are removed.
- **Categorical Encoding**: Nominal variables like 'Gender' and 'Department' are encoded using one-hot encoding.
- **Normalization**: Numerical features are scaled to the range [0, 1] using min-max normalization.
- **Handling Class Imbalance**: The Adaptive Synthetic Sampling (ADASYN) technique is applied to balance the dataset by generating synthetic samples for the minority class .

Data Preprocessing

1) Data description: The initial step in carrying out this study is performing a data pre-preprocessing task. This study produces a data quality report to detect outliers and any unusual pattern about the dataset using statistical methods. Tables II and III show the data quality report of the dataset.

2) Detecting outliers: In addition to the above data quality report, forty-five (45) outliers were detected using the Interquartile Range filter based on the initial raw dataset, and the outliers were then checked. Those findings require further preprocessing, which are data cleaning, data reduction, and data transformation. There are also no missing values that are in existence, and the given data is complete.

3) Data visualization: An overview to understand each attribute pattern should be carried out and examined through data visualization. From the data visualization, we can see that a few attributes need to be examined to ensure accuracy during the model classification process. Fig. 1 shows the data visualization of each attribute in the dataset.

4) Data cleaning and reduction: The dataset is considered high dimensional as it consists of 35 attributes. Any irrelevant attributes that are not contributing to the objectives of this study should be removed. Based on the data quality report in Table III and data visualization in Fig. 1, „EmployeeCount," „StandardHours" and „Over18" features can be removed in view that the cardinality/distinction is „1", which means it has the same values throughout the data. Other than that, „EmployeeNumber" is found not useful for the modeling and prediction process and can be removed from the dataset. No spelling inconsistencies were detected as inconsistencies may cause problems in later merges or transformations. Further description of data cleaning and reduction is explained in Table IV.

5) Normalization and discretization: During the data transformation in the preprocessing stage, feature scaling or normalization is applied. Normalization is a method used to standardize the range of independent variables or features of data [23]. Applying feature scaling or normalization can avoid dependency on the choice of measurement units on attributes. This process made the range of features of data fall between 0 and 1. The data cleaning and reduction were performed, which include the discretization process and change of attribute type from numerical to nominal. Four (4) attributes were removed based on the findings above, leaving the remaining 30 attributes. No outliers were detected after the interquartile filter was regenerated.

TABLE 2 THE DATA QUALITY REPORT (CATEGORICAL ATTRIBUTES)

No.	Feature Name	Count	% of Missing Value	Card.	Mode	Mode Freq.	Mode %	2nd Mode	2nd Mode Freq.	2nd Mode %
1	Attrition	1470	0	2	No	1233	84	Yes	237	16
2	BusinessTravel	1470	0	3	Travel Rarely	1043	71	Travel Frequently	277	19
3	Department	1470	0	3	R & D	961	65	Sales	446	30
4	Education	1470	0	5	3	473	32	4	340	23
5	Education Field	1470	0	6	Life Science	606	41	Medical	464	32
6	Environment Satisfaction	1470	0	4	3	453	31	4	446	30
7	Gender	1470	0	2	Male	882	60	Female	588	40
8	Job Involvement	1470	0	4	3	868	59	2	375	25
9	Job Level	1470	0	5	1	543	37	2	534	36
10	Job Role	1470	0	9	Sales Exec	326	22	Research Scientist	292	20
11	Job Satisfaction	1470	0	4	4	459	31	3	442	30
12	Marital Status	1470	0	3	Married	673	46	Single	470	32
13	Over 18	1470	0	1	Y	1470	100	-	-	-
14	Over Time	1470	0	2	No	1054	72	Yes	416	28
15	Performance Rating	1470	0	2	3	1244	85	4	226	15
16	Relationship Satisfaction	1470	0	4	3	459	31	4	432	29
17	Stock Option Level	1470	0	4	0	631	43	1	596	41
18	Work Life Balance	1470	0	4	3	893	61	2	344	23

TABLE 3 THE DATA QUALITY REPORT (CONTINUOUS ATTRIBUTES)

No	Feature Name	Count	% of Missing Value	Cardinality	Min	1st Qrt	Mean	Medi an	3rd Qrt	Max	Std. Dev
1	Age	1470	0	43	18.00	30.00	36.92	36.00	43.00	60.00	9.14
2	DailyRate	1470	0	886	102.00	465.00	802.49	802.00	1157.00	1499.00	403.50 9
3	DistanceFromHome	1470	0	29	1.00	2.00	9.19	7.00	14.00	29.00	8.11
4	Employee Count	1470	0	1	1.00	1.00	1.00	1.00	1.00	1.00	0.00
5	Employee Number	1470	0	1470	1.00	491.25	1024.86 5	1020. 5	1556.00	2068.00	602.02 4
6	Hourly Rate	1470	0	71	30.00	48.00	65.89	66.00	83.00	100.00	20.33
7	MonthlyIncome	1470	0	1349	1009.00	2911.00	6502.93	4919.00	8380.00	19999.00	4707.96
8	Monthly Rate	1470	0	1427	2094.00	8047.00	14313.1 03	14235.50	20462.00	26999.00	7117.79
9	NumCompaniesWorked	1470	0	10	0.00	1.00	2.69	2.00	4.00	9.00	2.50
10	PercentSal ryHike	1470	0	15	11.00	12.00	15.21	14.00	18.00	25.00	3.66
11	StandardH ours	1470	0	1	80.00	80.00	80.00	80.00	80.00	80.00	0.00
12	TotalWorkingYears	1470	0	40	0.00	6.00	11.28	10.00	15.00	40.00	7.78
13	TrainingTimesLastYear	1470	0	7	0.00	2.00	2.80	3.00	3.00	6.00	1.29
14	YearsAtCompany	1470	0	37	0.00	3.00	7.01	5.00	9.00	40.00	6.13
15	YearsInCurrentRole	1470	0	19	0.00	2.00	4.23	3.00	7.00	18.00	3.62
16	YearsSinceLastPromotion	1470	0	16	0.00	0.00	2.19	1.00	3.00	15.00	3.22
17	YearsWithCurrManager	1470	0	18	0.00	2.00	4.12	3.00	7.00	17.00	3.57

TABLE 4. THE DATA QUALITY REPORT (CATEGORICAL ATTRIBUTES)

No.	Feature Name	Type of Data	Type of Data	Data Description	Pre-processing action/Findings
1	Age	Continuous	Numeric	The age of individual employee	Min = 18, max = 60 Normalize, Discretize
2	Attrition	Categorical	Nominal	Employee leaving the company (Yes, No)	Set to class
3	BusinessTravel	Categorical	Nominal	Business travel frequency (No Travel, Travel Frequently, Travel Rarely)	Retain
4	DailyRate	Continuous	Numeric	Salary Level	Normalize, Discretize
5	Department	Nominal	Nominal	Employee department (HR, R&D, Sales)	Retain
6	DistanceFromHome	Continuous	Numeric	The distance from work to home	Min = 1, Max = 29 Normalize, Discretize
7	Education	Categorical	Numeric	Level of education attained (1 = 'Below Collage', 2 = 'College', 3 = 'Bachelor', 4 = 'Master', 5 = 'Doctor')	Change to Nominal
8	EducationField	Nominal	Nominal	Field of education (HR, Life Sciences, Marketing, Medical Sciences, Others, Technical)	Retain
9	EmployeeCount	Continuous	Numeric	Count of instance	Cardinality = 1 - To remove
10	EmployeeNumber	Continuous	Numeric	Employee ID	Cardinality = 1470 - To remove
11	EnvironmentSatisfaction	Categorical	Numeric	Employee satisfaction with the environment (1 = 'Low', 2 = 'Medium', 3 = 'High', 4 = 'Very High')	Change to Nominal
12	Gender	Categorical	Nominal	Female, Male)	Retain
13	HourlyRate	Continuous	Numeric	Hourly Salary	Normalize, Discretize
14	JobInvolvem ent	Categorical	Numeric	Job Involvement (1 = 'Low', 2 = 'Medium', 3 ='High', 4 = 'Very High')	Change to Nominal
15	JobLevel	Categorical	Numeric	Level Of Job (1 to 5)	Change to Nominal

FIG. 1, EMPLOYEECOUNT

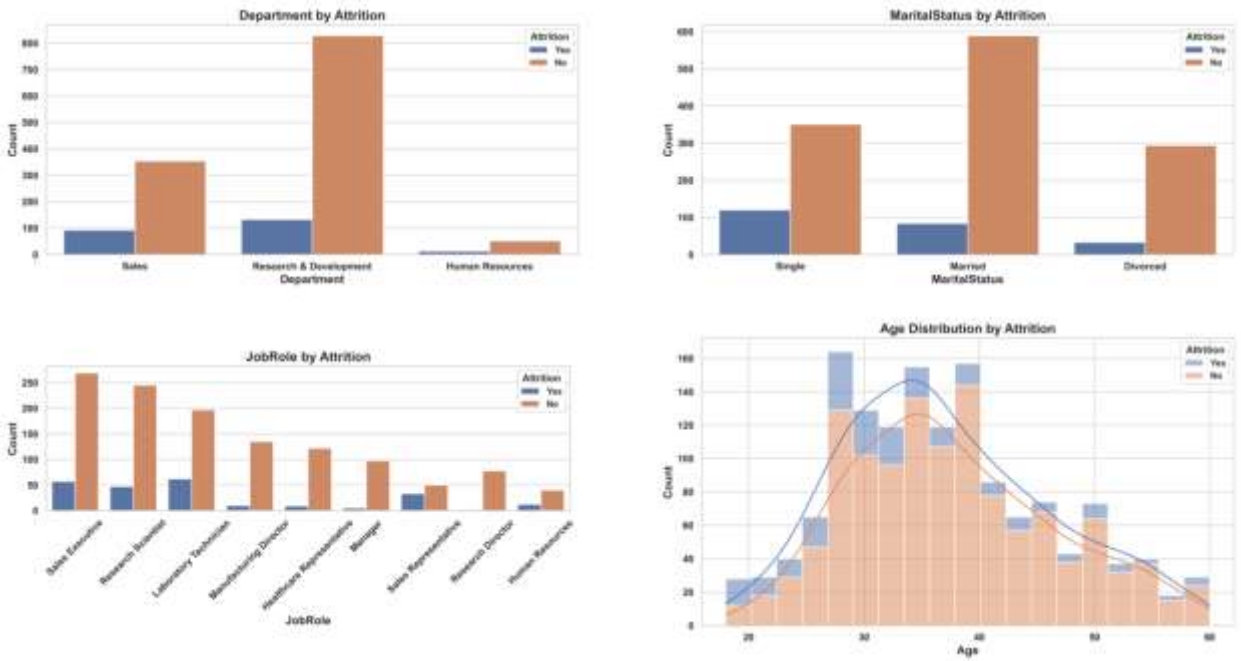


Figure 5. The pair plot data distributions analysis of features monthly income, job level, total working hours, years at the company, and years with the current manager role.

The classification of Employee Attrition Artificial Neural Network Classification Method

Figure 2 depicts the flowchart for the Artificial Neural Network classification method. In this contribution, the proposed ANN design used for classifying the roads into three categories.

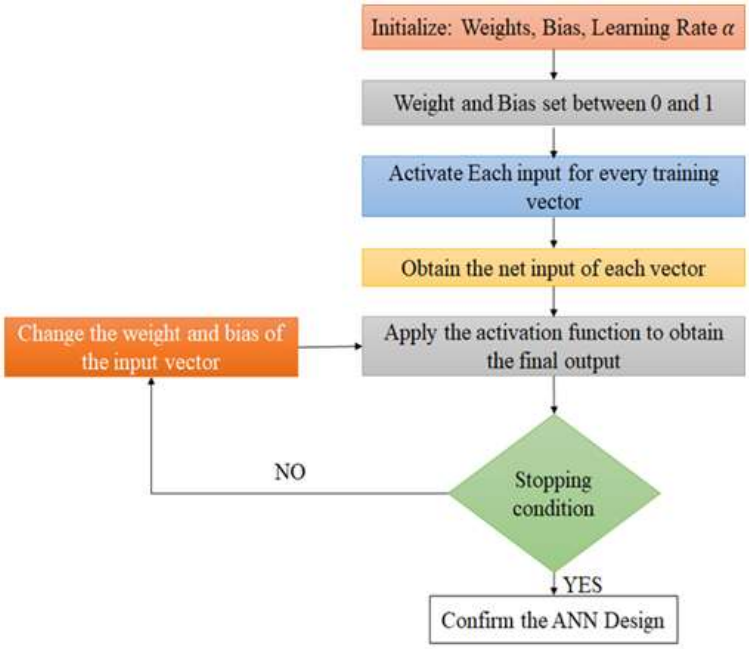


Figure 2: Flowchart for the Classification of Employee attritionusing Artificial Neural Network

In this contribution, ANN is used for classification of employee attritionin the employee attritiontransportation domain. In order to form a feed-forward Artificial Neural Network(a Multilayer perceptron(MLP)), the collection of non-linear neurons is connected to one another.This technique is known to be very useful for prediction and classification issues. Cross-validation is used to determine the 'optimal' number of hidden layers and neurons which were relied on the experimental design of the employee attrition.

Particularly, the training of the MLP began from a small number of neurons, and with only one hidden layer, which measures the error ratio of the trained BP on holdout samples, slowly increasing the number of neurons at the hidden layer in which the performance of the trained phase on holdout samples has begun to go down due to the problem of overtraining. Thus, we obtain the best number of neurons for the hidden layer of the ANN.The weight of the ANN can be updated by using Particle Swarm Optimization algorithm.

Pseudocode for proposed ANN Design

Step 1: Initialize the following to start the training of MLP Neural Network
Weights
Bias
Learning Rate α
For easy calculation and simplicity, weights and bias must be set equal to 0 and the learning rate must be set equal to 1.
Step 2: Continue step 3-8 when the stopping condition is not true.
Step 3: Continue step 4-6 for every training vector x.
Step 4: Activate each input as follows:

$$x_i = s_i (i = 1 \text{ to } n)$$

Step 5: Obtain the net input with the following relation:

$$y_{in} = b + \sum_{i}^{n} x_i w_{ij}$$

Step 6: Apply the following activation function to obtain the final output for the each input unit j=1 to m.

$$f(y_{in}) = \begin{cases} 1, & \text{iff} y_{inj} > \theta \\ 0, & \text{if } -\theta \le y_{inj} \le \theta \\ -1, & \text{if } y_{inj} < -\theta \end{cases}$$

Step 7: Adjust the weight and bias for x=1 to n and j=1 to m as follows using Particle Swarm Optimization:
Case 1: if $y_j \neq t_j$ then

$$w_{ij}(new) = w_{ij}(old) + \propto t_j x_j$$
$$b_j(new) = b_j(new) + \propto t_j$$

Case 2: if $y_j = t_j$ then

$$w_{ij}(new) = w_{ij}(old)$$
$$b_j(new) = b_j(old)$$

Here 'y' is the actual output and 't' is the desired /target output.

Step 8: Test for the stopping criteria which will happen when there is no change in the weight

3.3. Feature Selection

Chi-square (χ^2) tests are conducted to identify the most influential features affecting attrition. Key features include 'OverTime', 'JobLevel', and 'MonthlyIncome' .

3.4. Model Architecture

The Enhanced Deep Classifier comprises:

- **Input Layer**: 53 neurons corresponding to the number of features.
- **Hidden Layers**: Seven layers with 100 neurons each, utilizing the Softplus activation function to capture complex relationships.
- **Output Layer**: One neuron with a sigmoid activation function for binary classification.
- **Loss Function**: Binary cross-entropy.
- **Optimizer**: Adam optimizer with an initial learning rate of 0.01.

3.5. Model Training and Validation

The dataset is split into training (70%) and testing (30%) subsets. K-fold cross-validation is employed to assess model performance and mitigate over fitting .

4. Results and Discussion

The Enhanced Deep Classifier achieves an accuracy of 92%, precision of 0.90, recall of 0.93, and F1-score of 0.91. These results surpass those of traditional machine learning models and demonstrate the efficacy of deep learning in predicting employee attrition.

5. Conclusion

The Enhanced Deep Classifier offers a robust framework for predicting employee attrition, providing organizations with valuable insights to implement effective retention strategies. Future work will explore the integration of natural language processing techniques to analyze textual data, further enhancing prediction accuracy.

References

1. Alduayj, M., & Rajpoot, N. (2020). Employee Attrition Prediction Using Deep Neural Networks. MDPI. FTST Journal. (2024). Machine Learning Model for Employee Attrition Prediction.
2. Li, W. (2023). A Transformer-Based Deep Learning Framework to Predict Employee Attrition. PeerJ Computer Science.
3. Pratibha, G., & Hegde, N. P. (2023). Employee Attrition Prediction Based on Grey Wolf Optimization and Deep Neural Networks. International Journal on Recent and Innovation Trends in Computing and Communication, 11(10), 1338–1345.
4. University of Toronto Machine Intelligence Team. (2024). Employee Attrition Factors Prediction. Medium.
5. Gandomi, A.H.; Chen, F.; Abualigah, L. Machine Learning Technologies for Big Data Analytics. Electronics 2022, 11, 421.
6. Jia, X.; Cao, Y.; O'Connor, D.; Zhu, J.; Tsang, D.C.W.; Zou, B.; Hou, D. Mapping soil pollution by using drone image recognition and machine learning at an arsenic-contaminated agricultural field. Environ. Pollut. 2021, 270, 116281.
7. Reshma Ramchandra, N.; Rajabhushanam, C. Machine learning algorithms performance evaluation in traffic flow prediction.Mater. Today Proc. 2022, 51, 1046–1050.
8. Aljedani, N.; Alotaibi, R.; Taileb, M. HMATC: Hierarchical multi-label Arabic text classification model using machine learning. Egypt. Inform. J. 2021, 22, 225–237.
9. Aggarwal, S.; Singh, M.; Chauhan, S.; Sharma, M.; Jain, D. Employee Attrition Prediction Using Machine Learning Comparative Study. Smart Innov. Syst. Technol. 2022, 265, 453–466.
10. Jadhav, A.; Mostafa, S.M.; Elmannai, H.; Karim, F.K. An Empirical Assessment of Performance of Data Balancing Techniques in Classification Task. Appl. Sci. 2022, 12, 3928.
11. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.E.; Arshad, H. State-of-the-art in artificial neural networkapplications: A survey. Heliyon 2018, 4, e00938.
12. 13. Ganthi, L.S.; Nallapaneni, Y.; Perumalsamy, D.; Mahalingam, K. Employee Attrition Prediction Using Machine Learning Algorithms. Lect. Notes Netw. Syst. 2022, 288, 577–596.
13. Jiang, Z.; Gao, B.; He, Y.; Han, Y.; Doyle, P.; Zhu, Q. Text Classification Using Novel TermWeighting Scheme-Based ImprovedTF-IDF for Internet Media Reports. Math. Probl. Eng. 2021, 2021, 6619088.
14. Qutub, A.; Al-Mehmadi, A.; Al-Hssan, M.; Aljohani, R.; Alghamdi, H.S. Prediction of Employee Attrition Using Machine Learning and Ensemble Methods. Int. J. Mach. Learn. Comput. 2021, 11, 110–114.
15. Habous, A.; Nfaoui, E.H.; Oubenaalla, Y. Predicting Employee Attrition using Supervised Learning Classification Models. In Proceedings of the 2021 Fifth International Conference on Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, 20–22 October 2021.
16. Najafi-Zangeneh, S.; Shams-Gharneh, N.; Arjomandi-Nezhad, A.; Zolfani, S.H. An Improved Machine Learning-Based Employees Attrition Prediction Framework with Emphasis on Feature Selection. Mathematics 2021, 9, 1226.
17. Pratt, M.; Boudhane, M.; Cakula, S. Employee attrition estimation using random forest algorithm. Balt. J. Mod. Comput. 2021, 9, 49–66.
18. Sadana, P.; Munnuru, D. Machine Learning Model to Predict Work Force Attrition. In Proceedings of the 2021 6th International Conference for Convergence in Technology (I2CT), Pune, India, 2–4 April 2021.
19. Kaya, ˙I.E.; Korkmaz, O. Machine Learning Approach for Predicting Employee Attrition and Factors Leading to Attrition. Cukurova Univ. J. Fac. Eng. 2021, 36, 913–928.
20. Mazumder, R.K.; Salman, A.M.; Li, Y. Failure risk analysis of pipelines using data-driven machine learning algorithms. Struct. Saf. 2021, 89, 102047.

21. Mate, Y.; Potdar, A.; Priya, R.L. Ensemble Methods with Bidirectional Feature Elimination for Prediction and Analysis of Employee Attrition Rate During COVID-19 Pandemic. Lect. Notes Electr. Eng. 2022, 806, 89–101.
22. Joseph, R.; Udupa, S.; Jangale, S.; Kotkar, K.; Pawar, P. Employee attrition using machine learning and depression analysis. In Proceedings of the 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 6–8 May 2021; pp. 1000–1005.