International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *18 August 2023* ***Revised:*** *08 October 2023* ***Accepted:*** *22 October 2023*

Employee Attrition Prediction based on Grey Wolf Optimization and Deep Neural Networks

[1]G Pratibha, [2]Nagaratna P. Hegde
[1] Research Scholar, Department of Computer Science and Engineering, JNTU, Hyderabad, Telangana, India
pratibhaphd123@gmail.com
[2] Professor**,** Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, Telangana,India

Abstract: Despite the constructive application of promising technologies such as Neural Networks, their potential for predicting human resource management outcomes still needs to be explored. Therefore, the primary aim of this paper is to utilize neural networks and meta-heuristic technologies to predict employee attrition, thereby enhancing prediction model performance. The conventional Grey Wolf optimization optimization (GWO) has gained substantial attention notice because of its attributes of robust convergence, minimal parameters, and simple implementaton. However, it encounter problems with slow convergence rates and susceptibility to local optima in practical optimization scenarios. To address these problems, this paper introduces an enhanced Grey Wolf Optimization algorithm incorporating the utilization of Cauchy-Gaussian mutation, which contributes to enhancing diversity within the leader wolf population and enhances the algorithm's global search capabilities. Additionally, this work preserves exceptional grey wolf individuals through a greedy selection of 2 mechanisms to ensure accelerated convergence. Moreover, an enhanced exploration strategy is suggested to expand the optimization possibilities of the algorithm and improve its convergence speed. The results shows that the proposed model achieved the accuarcy of 97.85%, precision of 98.45%, recall of 98.14%, and f1-score of 97.11%. Nevertheless, this paper extends its scope beyond merely predicting employee attrition probability and activities to enhance the precision of such predictions by constructing an improved model employing a Deep Neural Network (DNN)..

Keywords- Machine learning; Employee attrition, prediction model, Grey Wolf Optimization, Cauchy-Gaussian mutation

I. INTRODUCTION

In the contemporary competitive economy characterized by increasing technological specialization, the acquisition, study and analysis of data have given rise to a new concept known as the "knowledge economy." Information technologies not solely function as a data source but also assume a vital role in facilitating data analysis. This enable the processing of large data sets and the extraction of valuable information from them. Data has become a strategic asset for a wide range of organizations across various sectors, including those involved in business processes. The adoption of new technologies benefits all types of organizations [1], and the collection, management, and analysis of data offer numerous advantages in terms of efficiency and gaining a competitive edge. Certainly, examining extensive datasets can result in improvements in decision-making, the attainment of established corporate objectives, and heightened business competitiveness [2, 3]. Predicting employee attrition in advance can aid in preventing or minimizing its repercussions on an organization. Some literature suggests that content and motivated employees tend to exhibit greater creativity, productivity, and overall performance [4].

Organizations can leverage their HR data to make such predictions by employing predictive models designed for this purpose. In recent times, artificial intelligence (AI) has been applied across diverse domains, such as healthcare, education, the economy, and governance [5, 6]. Notably, there has been a growing focus on using AI for predicting employee attrition, reflecting the increasing research attention in this area. Furthermore, the growing volume of data on this subject has spurred an increased number of studies in this field [7, 8]. This research paper specifically centers on forecasting employee attrition using deep neural networks, employing the IBM Watson dataset for training and testing purposes. This dataset comprises 35 features for a total of 1470 samples, categorized into two classes: current employees and former employees. Notably, the samples are imbalanced, consisting of 237 positive instances (former employees) and 1233 negative instances (current employees). This class imbalance presents a considerable challenge in the prediction process. Workplace absenteeism may be predicted using a deep neural network, as shown in [9]. They demonstrated that their technique was 97.5% accurate compared to the 85% accuracy of the standard model. Later, in [10], an enhanced Deep Belief Network based on GWO is used to predicting employee turnover. Slow convergence time and a propensity to settle on a suboptimal solution are two issues that plague GWO.

To solve this the present paper contributions can be concisely summarized as follows: First, we have connected deep learning techniques along with preprocessing steps to enhance the accuracy of employee attrition predictions. Second, we have

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 18 August 2023 **Revised:** 08 October 2023 **Accepted:** 22 October 2023*

conducted an in-depth analysis of dataset features to uncover their interrelationships and detect the most significant features using Cauchy-Gaussian mutation GWO. In order to guarantee the convergence speed of the GWO, the present study keeps grey wolf individuals using the greedy selection process. To increase the convergence rate of the algorithm's accuracy and search space, a new search technique was devised. Third, for an inclusive assessment of the model's performance, we have tested it on both balanced and imbalanced datasets. Fourth, in contrast to several prior approaches, we have employed cross-validation to strictly evaluate the effectiveness of the work.

II. RELATED WORK

Researchers have explored the topic of employee attrition from various approaches. Some studies have delved into employee behavior to reveal the factors influencing their decisions to remain with or leave from an organization [11, 12].

In [13], the authors introduced a people analytics method for predicting employee outcomes attrition that shifts from a focus on big data to deep data, emphasizing data quality over quantity. This deep data-driven approach involves a mixed methodology to construct a relevant employee attrition model, identifying key employee attributes that impact attrition. The process began with a comprehensive collection of common features from the literature (an exploratory research phase). Afterward, a more targeted identification of essential attributes was conducted using survey and feature selection algorithms (a quantitative approach). Furthermore, this approach to predicting attrition depends on machine learning, deep learning, and ensemble models, which underwent testing on both extensive and moderately-sized simulated HR datasets, along with a genuine, smaller dataset comprising 450 responses. While rewards and compensation are typically considered crucial for retention, this approach explores a more nuanced perspective.

In [14], the authors' objective is to reveal the reasons behind employee attrition and establish a predictive model for attrition. They concentrated on examining organizational factors that contribute to attrition and tested four machine-learning techniques for comparison. Their refined approach using the Extra Trees Classifier (ETC) achieved an impressive accuracy rate of 93% in predicting employee attrition, outperforming recent state-of-the-art research. To identify key factors in attrition, they employed Employee Exploratory Data Analysis (EEDA), and their findings underscored the importance of monthly income, hourly rate, job level, and age as significant contributors.

In [15], the authors used the IBM analytics dataset, which contains 35 attributes for 1470 employees, and conducted preprocessing to remove irrelevant features. They created a balanced dataset from the original to ensure realistic results, evaluated various models using appropriate metrics, applied hyperparameter optimization to enhance performance, and employed cross-validation for an exhaustive assessment. The XG Boost classifier emerged as the most accurate, achieving an 85.5% accuracy rate, followed by the Random Forest classifier at 84%. Precision was also higher for the XG Boost classifier (0.72) compared to the Random Forest classifier (0.67).

In [16], the authors investigated the potential of data clustering methods in assessing worker performance and guiding management decisions. They looked at a wide range of performance indicators, such as word, timeliness, and articulacy. Predictions about staff incentives, designation changes, and terminations were among the article's outcomes. This approach aids in classifying inefficient employees, quantifying their inadequacies, and offers a straightforward framework for addressing these issues. They highlighted the serious role of employees in an organization's success. They suggested the use of a hybrid method combining Data Clustering and Decision Trees for predicting employee performance in the upcoming year.

In [17], the authors set out to forecast employee attrition within a company using the logistic regression method. They believe that machine learning is a valuable tool for predicting employee attrition because it operates without the biases that can result from human involvement. Moreover, the human resources department within a company requires a clear understanding of the most influential factors contributing to employee attrition. To address this, they introduced features to pinpoint influential factors and optimize the data training process, and the authors utilized various feature selection techniques. Their methodology revolved around predicting employee attrition, and they applied three distinct feature selection methods: information gain, select k-best, and recursive feature elimination (RFE). They further evaluated a 10-fold cross-validation procedure. In the context of predicting employee attrition, when employing the logistic regression method without employing feature selection, they achieved an accuracy rate of 0.865 and an AUC score of 0.932. Nevertheless, upon incorporating the RFE (Recursive Feature Elimination) feature selection method, they obtained the most favorable evaluation outcomes in comparison to information gain and select k-best methods, resulting in an accuracy rate of 0.853 and an AUC score of 0.925.

Despite progress made by previous solutions, there remains room for improving prediction accuracy to enhance confidence in the predictions. To address this, the proposed work incorporates deep learning techniques and data preprocessing methods with the goal of achieving higher prediction accuracy.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 18 August 2023 **Revised:** 08 October 2023 **Accepted:** 22 October 2023*

III. THE PROPOSED WORK

The proposed approach uses an attrition dataset sourced from the Kaggle data repository. This dataset comprises 34 attributes that provide information about various aspects of employee features. To minimize any potential errors affecting classifier accuracy, the data underwent preprocessing using the min-max normalization method. In this paper, a deep belief network employing a generative graphical model is employed. To further enhance the models accuracy, a meta-heuristic optimization technique known as grey wolf optimization is applied as shown in Figure 1.

Dataset description

The dataset utilized in this study originates from IBM Analytics [18], encompassing 35 distinct features related to 1470 employees. The attributes within the dataset, including their respective data types, are presented in Table 1. Of particular significance is the "Attrition" feature, which signifies the employee's decision, denoted as "Yes" (indicating departure from the company) or "No" (indicating retention within the company)..

Data preprocessing

The attrition dataset encompasses a diverse set of attributes, each with varying value ranges. When these attributes with extensive value ranges are directly input into a classifier, it can significantly impact the accuracy of the results. The utilization of these raw values in subsequent algorithmic processing can lead to substantial distortions in accuracy. To mitigate the effects of this disparity, Min-max normalization is applied. This normalization process ensures that all attributes are transformed to a uniform range spanning from 0 to 1. Prior to network processing, the data undergoes standardization through a linear mapping relationship.

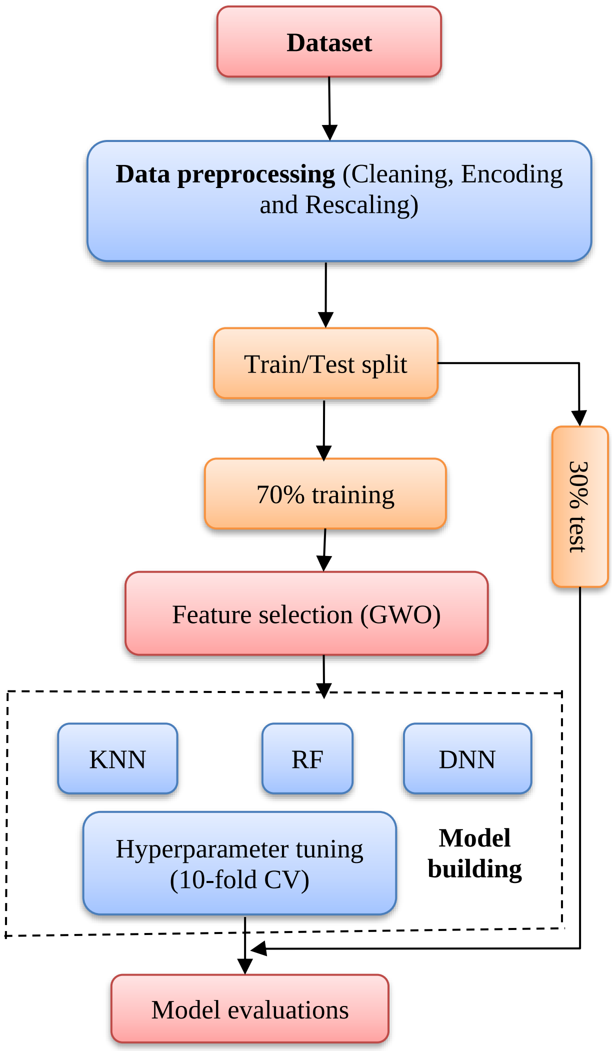


Figure 1: Proposed framework

Table 1. IBM dataset features

Feature Name	Type	Feature Name	Type
Age	Number	MonthlyIncome	Number
BusinessTravel	Category	MonthlyRate	Number
DailyRate	Number	NumCompaniesWorked	Number
Department	Category	Over18	Category
DistanceFromHome	Number	OverTime	Category
Education	Category	PercentSalaryHike	Number
EducationField	Category	PerformanceRating	Number
EmployeeCount	Number	RelationshipSatisfaction	Category
EmployeeNumber	Number	StandardHours	Number
EnvironmentSatisfaction	Category	StockOptionLevel	Category
Gender	Category	TotalWorkingYears	Number
HourlyRate	Number	TrainingTimesLastYear	Number
JobInvolvement	Category	WorkLifeBalance	Category
JobLevel	Category	YearsAtCompa	Number
EducationField	Category	YearsInCurrentRole	Number
JobRole	Category	YearsSinceLastPromotion	Number
JobSatisfaction	Category	YearsWithCurrentManager	Number
MaritalStatus	Category	Attrition	Category

The Min-Max normalization is mathematically represented as follows.

$$att_i new = \frac{att_i - min(att_i)}{max(att_i) - min(att_i)} \qquad (1)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 18 August 2023 **Revised:** 08 October 2023 **Accepted:** 22 October 2023*

In this representation, "att_i" represents the original attribute value, while "min" and "max" represent the minimum and maximum values found within the entire record of the respective attribute.

Classical Grey Wolf optimizer

The social structure within the grey wolf community is categorized into four distinct levels, as illustrated in Figure 2.

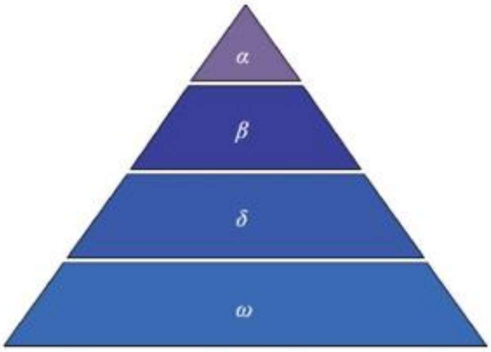


Figure 2: Grey Wolf social hierarchy

The hierarchical structure of grey wolves in the algorithm consists of four levels. The first level is referred to as the "α wolf," which serves as the decision maker among the wolves. This wolf possesses managerial abilities and corresponds to the optimal solution in the Grey Wolf Optimization (GWO) process. The second level, referred to as the "β wolf," and the third level, known as the "δ wolf," are designations for sub-optimal and third-optimal solutions within the GWO. Their primary function is to aid the α wolf in decision-making, collectively guiding and supporting other wolves in their pursuit of targets. The fourth level includes the "ω wolf," representing additional solutions in the optimization process. This level adjusts its position based on the decisions made by the α, β, and δ wolves. Throughout the algorithm's iterations, grey wolves at all levels engage in competitive interactions. Following each iteration, leader wolves must be reselected based on the distances between individual grey wolves and the target prey. The updating of the positions of the α, β, and δ wolves within the grey wolf population is contingent on the prey's position, as expressed in formula (2):

$$X(t+1) = X_p(t) - A.D \tag{2}$$

Here, $X(t + 1)$ represents the updated position of a grey wolf, $X_p(t)$ denotes the prey's position "A" represents a vector of coefficients, while "D" represents the distance between a grey wolf and its prey, calculated as outlined in formula (3):

$$D = \left|C.X_p(t) - X(t)\right| \tag{3}$$

The coefficient vector C is determined by formula (4):

$$C = 2, r_1 \tag{4}$$

The coefficient vector A is computed using formula (5):

$$A = 2a.r_2 - a \tag{5}$$

In these formulas, $X(t)$ represents the current position of a grey wolf, the linear decrease occurs from 2 to 0 as the iterations progress, and r_1 and r_2 represent random vectors within the range of [0, 1]. According to the grey wolf social hierarchy, ω wolves rely on leader wolves for position adjustments. To calculate the separation of ω wolves and each leader wolf, use equation (6). Afterward, the movement direction for each individual grey wolf is determined based on the formulas (7) and (8).

$$\begin{cases} D_\alpha = |C_1.X_\alpha - X| \\ D_\beta = |C_2.X_\beta - X| \\ D_\delta = |C_1.X_\delta - X| \end{cases} \tag{6}$$

$$\begin{cases} X_1 = X_\alpha - A_1 - D_\alpha \\ X_2 = X_\beta - A_2 - D_\beta \\ X_3 = X_\delta - A_3 - D_\delta \end{cases} \tag{7}$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{8}$$

In these equations, X_α, X_β, and X_δ represent the positions of the α, β, and δ wolves, respectively, while X denotes the current position of the individual grey wolf. D_α, D_β, and D_δ represent the distances between the individual grey wolf and each of the leader wolves. Lastly, $X(t + 1)$ signifies the grey wolf's updated position after the update process.

Proposed method

Cauchy-Gaussian mutation. The Cauchy-Gaussian mutation (CGM) strategy is introduced to address a limitation observed in the later iterations of the traditional GWO algorithm. In these later stages, grey wolves tend to converge toward the α wolf, resulting in reduced diversity in local searches within the population and premature convergence. To mitigate this issue, the incorporation of the Cauchy-Gaussian mutation operator is applied with the goal of augmenting diversity within the leader wolf group and enhancing their ability to perform localized searches. Following each iteration, the α, β, and δ wolves are selected for mutation, and a comparison is conducted between the mutated positions and the original positions, guided by a prioritized selection mechanism favoring individuals with superior fitness to advance to the subsequent iteration. The mathematical representation of the Cauchy-Gaussian mutation strategy is outlined as follows:

$$U_{leader}(t+1) = X_{leader}[1 + \lambda_1 cauchy(0, \sigma^2) + \lambda_1 Gauss(0, \sigma^2)] \tag{9}$$

$$\sigma = exp\left(\frac{f(X_{leader}) - f(X_\alpha)}{|fX_\alpha|}\right) \tag{10}$$

$$X(t+1) = \begin{cases} U_{leader}(t+1) f\big(U_{leader}(t+1) \le f(X_{leader})\big) \\ X_{leader}\ otherwise \end{cases} \tag{11}$$

In these equations, $U_{leader}(t + 1)$ signifies the position of leader wolves after mutation, while X_{leader} represents their current positions. The terms "Cauchy (0, σ^2)" and " Gauss (0, σ^2) " refer to random variables that follow the Cauchy distribution and Gaussian distribution, respectively. The word " $f(X_{leader})$" refers to the fitness value of leader wolves, whereas "$f(X_\alpha)$" indicates the fitness value of the α wolf.

$$\lambda_1 = 1 - \frac{t^2}{T^2} \quad (12)$$

$$\lambda_2 = \frac{t^2}{T^2} \quad (13)$$

Here, "t" denotes the ongoing iteration count, while "T" signifies total iteration count.

Improved search strategy: The classic GWO method has few search variables, thus rendering it straightforward to use. However, its global search capability is relatively weak, and it often tends to get trapped in local optima in certain cases. To address this limitation, this study introduces an enhanced search strategy aimed at bolstering the algorithm's global search ability and broadening the exploration space. This improved search strategy expands the global search space while preserving the existing position of a grey wolf individual, represented as $X(t)$, created using the conventional GWO position update formula (8). The mathematical expression for this improved search approach is as follows

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1 \times |X_{rand}(t) - 2 \times r_2 \times| \\ X(t), r_3 \geq 0.5\ X_q(t) - X_{avg}(t) - r_3 \\ \times (lb + r_4 \times (ub - lb), r_3 < 0.5) \end{cases} \quad (14)$$

$U(t+1)$ is the position of a grey wolf after applying the improved search strategy, $X_{rand}(t)$ is the position of a randomly selected grey wolf from the population at the tth iteration, $X(t)$ is the current position, and "r_1, r_2, r_3, r_4, r_5" are random vectors within [0,1]. In addition, $X_\alpha(t)$ represents the current location of the α wolf, $X_{avg}(t)$ reflects the grey wolf population average, and ub and lb indicate decision variable upper and lower limits.

$$X(t+1) = \begin{cases} U(t+1), f\big(U(t+1)\big) \leq f\big(X(t)\big) \\ X(t), otherwise \end{cases} \quad (15)$$

In equation (15), $X(t+1)$ is the location of a grey wolf in the $(t+1)$th iteration, $f(U(t+1))$ is the fitness value after adjusting the position employing the better search method, and $f(X(t))$ is the current fitness value. This enhanced search strategy generates new solutions in proximity to random solutions or optimal solutions during the iteration, facilitating improved search and communication among grey wolf individuals. If the exploration defined by formula (14) fails to yield a superior position, the traditional GWO method is employed to update the grey wolf individual's position.

Classification model:

The modeling process involves the selection of machine learning models, including Random Forest (RF), K-nearest neighbor (KNN), and Deep Neural Network (DNN), as employed in the paper. DNN, or Deep Neural Network, can be described as a neural network with multiple hidden layers. Generally, the first layer serves as the input layer, the final layer as the output layer, and the intermediate layers as hidden layers. The training process of a DNN consists of two key steps:

Forward Propagation: This step involves using the output from the upper layer to calculate the output of the subsequent layer.

Back Propagation: The gradient descent algorithm is utilized to propagate calculation errors layer by layer, thereby adjusting the weights of each layer. This process optimizes the network model through iterative cycles of forward propagation and back propagation. Each hidden unit, denoted as "j," typically employs a specific activation function (in this paper, the sigmoid function is utilized). This function maps the total input received from the preceding layer, denoted as "x_j," to the output for the next layer, denoted as "y_j."

$$y_j = logistic(x_j) = \frac{1}{1 + e^{-x_j}} \quad (16)$$

$$x_j = b_j + \sum_i y_i \omega_j \quad (17)$$

In this notation, ω_j represents the weight on a link to unit j from unit i in the underlying layer, and b_j is the bias of unit **j.** In the context of multiclass classification, for an output unit labeled as "j" the softmax function is employed to transform its cumulative input, denoted as "x_j" into class probabilities represented by "p_j"

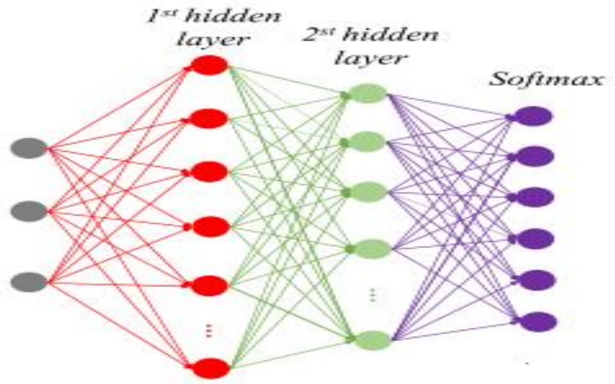


Figure 3: Structure of the proposed DNN model

Here, "k" serves as an index spanning all available classes.

$$p_j = \frac{\exp(x_j)}{\sum_k \exp(x_k)} \quad (18)$$

Figure 3 illustrates the schematic diagram of the network, where the input to the network consists of the extracted features, and the network produces 6 output signals.

IV. RESULTS AND DISCUSSIONS

In this experimental setup, we utilized an Intel (TM)-i5 processor running at a clock rate of 3.2GHz and equipped with 8GB of main memory. The feature extraction techniques were implemented on a system running Windows 7 Ultimate, with Matlab 2016 and Python 3.5 as the software environments. To assess the algorithm's performance, we calculated accuracy

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 18 August 2023 **Revised:** 08 October 2023 **Accepted:** 22 October 2023*

based on the percentage of successfully classified instances in the dataset. This evaluation criterion considers both true positives and true negatives, depending on the context. Additionally, we employed other performance evaluation metrics to provide a comprehensive assessment, including precision, recall, and F1 score. These metrics were employed to present the results of the experimental evaluations.

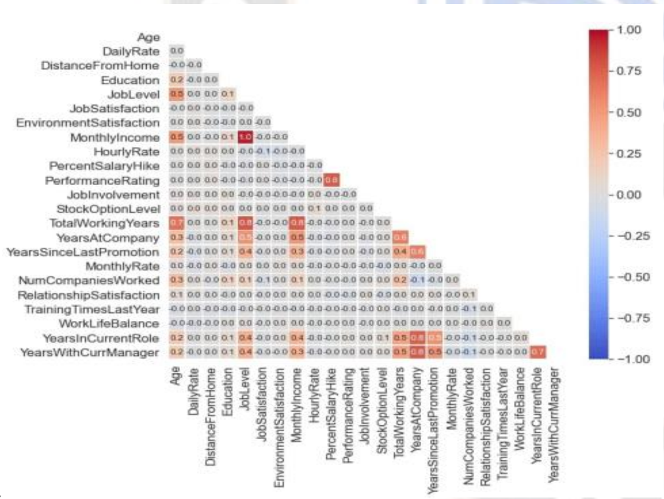
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \qquad (19)$$

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

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

$$\text{F1} - \text{score} = \frac{1}{\text{Precision}} + \frac{1}{\text{Recall}} \qquad (22)$$

In the context of these calculations, TP stands for true positives, FP stands for false positives, TN denotes true negatives, and FN signifies false negatives. Typically, a correlation matrix is employed to gain insights into the relationships among the features within a dataset. Figure 4 illustrates the correlation matrix of our dataset. The cell colors in the matrix range from blue to red. Grey cells indicate no correlation, while variations in red indicate a high level of correlation. Conversely, blue variations indicate a negative correlation among the dataset's features.



.

Figure 4: Correlation Matrix

The significance of individual features in making a forecast varies. Chi-square x2 ranking was used to prove our point. As can be seen in Figure 5, the most salient characteristics are OverTime, JobLevel, and MonthlyIncome.

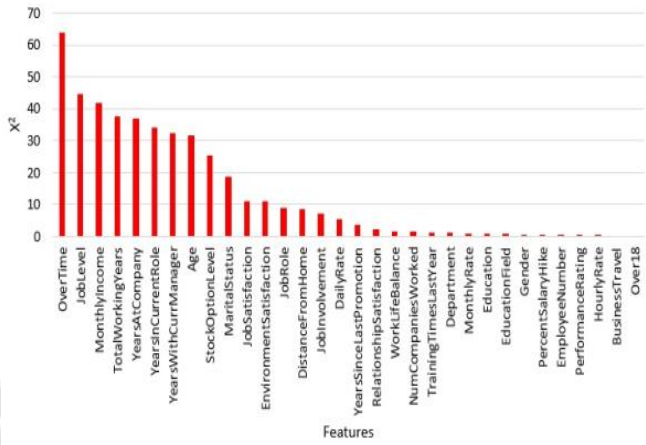


Figure 5: Important Features

We employed grey wolf optimization for feature selection, successfully identifying crucial features for our study. In this analysis, we considered two machine learning models, along with one Deep Neural Network, which has garnered significant attention from recent researchers. Specifically, we examined the following machine learning models: DNN, KNN, and RF. The investigation primarily revolved around key operational aspects, including the CPU processing time required for training different models, without the use of a GPU, as well as the model's memory size. Additionally, we conducted an evaluation of machine learning metrics for each model, encompassing Cross Validation Mean Score, Model Accuracy, F1-Score, Classification Precision, and Recall.

Table 2: Comparison of performance metrics for 3 models without feature selection

Type of Model	**Accuracy**	**Precision**	**Recall**	**F1-score**
KNN	90.12	89.12	88.14	88.77
RF	89.34	88.23	87.12	88.54
DNN (Our work)	96.78	97.23	97.44	95.66

From Figure 6 it is clear that the Proposed DNN is showing better performance as compared to all other two ML approaches.

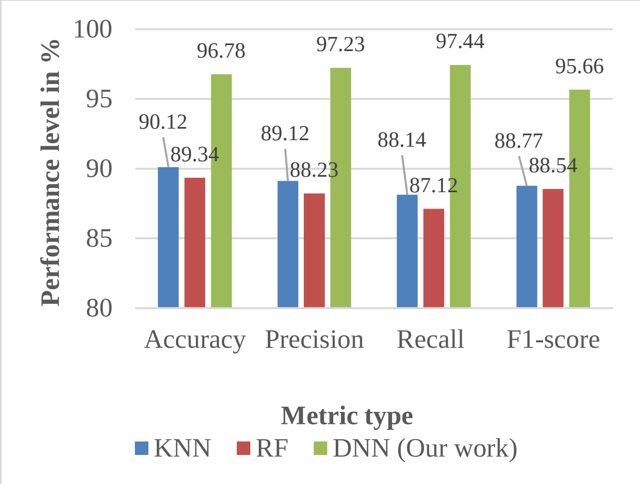


Figure 6: Comparison Performance metrics of 3 models

Table 3: Comparison of performance metrics for 3 models with feature selection (GWO)

Type of Model	Accuracy	Precision	Recall	F1-score
KNN	91.12	90.12	89.14	89.77
RF	90.16	89.43	88.45	89.65
DNN (Our work)	97.85	98.45	98.14	97.11

From Figure 7 it is clear that Our Proposed DNN with feature selection using GWO is showing better performance as compared to all other two ML approaches.

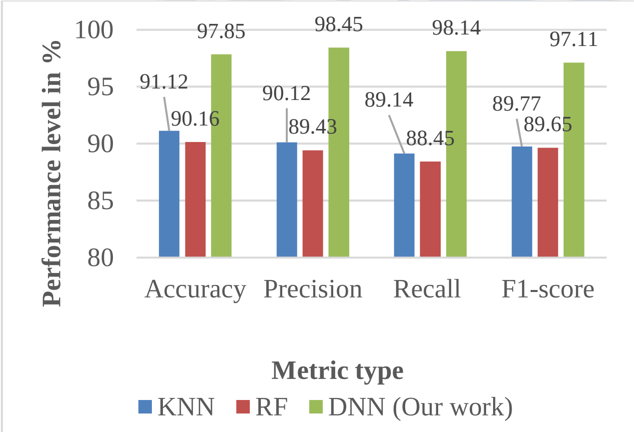


Figure 7: Comparison Performance metrics of 3 models (with feature selection)

Table 4: Computational time in sec

Type of Model	Base Model	Hyper Parameter Tuning
KNN	19.3	2.65
RF	28.93	2.97
DNN (Our work)	16.12	2.11

From Figure 8 it is clear that the proposed DNN is takes less time as compared to all other two ML approaches.

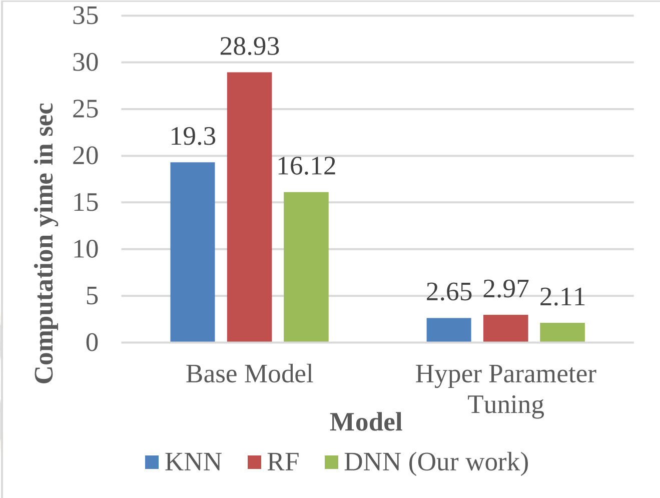


Figure 8: Comparison of computation time of 3 models

V. CONCLUSION

In this paper, the proposed work can assist the human resources department in gaining insights into employee attrition within the organization. This method leverages employee indications to predict the potential risk of attrition, providing valuable information to management. Additionally, we explore the correlations among various features within the dataset. To enhance the feature selection process, we introduce a wrapper feature selection method based on CGM Grey Wolf optimization and Deep Neural Networks (DNN). This method effectively identifies relevant features while eliminating redundant and irrelevant ones from the IBM HR dataset. The results indicate that the most significant factors impacting employee decisions are overtime hours, job level, and monthly income. We also address the challenge posed by the imbalanced nature of the IBM analytics dataset by creating a synthetic version that enables the building of a robust classifier capable of delivering realistic predictions. From the experiments, we achieved an accuracy of 97.85% through 10-fold cross-validation, superior to the performance of previously presented methods. However, the computation time is 2.11 sec. Future work might refine the findings by factoring in attractive job openings and the presence of negative working circumstances, both of which have a positive relation to employee attrition.

REFERENCES

[1] X. Mirjalili, S., Mirjalili, M. & Lewis, A. Grey Wolf optimizer. Adv. Eng. Softw. 69, 46–61 (2014).

[2] Kennedy, N., Eberhart, C. Particle swarm optimization. In Proceedings of ICNN'95—International Conference on Neural Networks, vol. 1–6 1942–1948 (2002).

[3] Mirjalili, S. & Lewis, A. The Whale Optimization algorithm. Adv. Eng. Softw. 95, 51–67 (2016).

[4] Xue, J. & Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. Syst. Sci. Control Eng. 8(1), 22–34 (2020).

[5] Salgotra, R., Singh, U. & Sharma, S. On the improvement in Grey Wolf Optimization. Neural Comput. Appl. 32, 3709–3748 (2020).
[6] Miao, D. et al. Parameter estimation of PEM fuel cells employing the hybrid Grey Wolf Optimization method. Energy 193, 571–582 (2020).
[7] Kulkarni, O. & Kulkarni, S. Process parameter optimization in WEDM by Grey Wolf optimizer. Mater. Today Proc. 5(2), 4402–4412 (2018).
[8] Luo, K. & Zhao, Q. A binary grey wolf optimizer for the multidimensional knapsack problem. Appl. Soft Comput. 83, 105645 (2019).
[9] Ali Shah, Syed & Uddin, M. Irfan & Aziz, Furqan & Ahmad, Shafiq & Al-Khasawneh, Mahmoud & Sharaf, Mohamed. (2020). An Enhanced Deep Neural Network for Predicting Workplace Absenteeism. Complexity. 2020. 1-12. 10.1155/2020/5843932.
[10] Usha, P.M. & Balaji, N.V.. (2021). A Grey Wolf Optimization Improved Deep Belief Network for Employee Attrition Prediction. International Journal of Engineering Trends and Technology. 69. 72-81. 10.14445/22315381/IJETT-V69I10P210.
[11] Zewen, L. et al. A hybrid grey wolf optimizer for solving the product knapsack problem. Int. J. Mach. Learn. Cybern. 12, 201–222 (2020).
[12] Kamboj, K., Bath, K. & Dhillon, S. Solution of non-convex economic load dispatch problem using Grey Wolf Optimizer. Neural Comput. Appl. 27(5), 1301–1316 (2015).
[13] Ben Yahia, Nesrine & Jihen, Hlel & Colomo-Palacios, Ricardo. (2021). From Big Data to Deep Data to Support People Analytics for Employee Attrition Prediction. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3074559.
[14] Raza, Ali. (2022). Predicting Employee Attrition Using Machine Learning Approaches. Applied Sciences. 12. 10.3390/app12136424.
[15] Bada, Olufunke & John Lekan, Akinode. (2022). Employee attrition prediction using machine learning algorithms.
[16] Sarker, Ananya & Shamim, S. M. & Zaman, Md. Shahid & Rahman, Md. (2018). Employee's Performance Analysis and Prediction Using K-means Clustering & Decision Tree Algorithm. 18.
[17] Wardhani, Fitri & Lhaksmana, Kemas. (2022). Predicting Employee Attrition Using Logistic Regression With Feature Selection. Sinkron. 7. 2214-2222. 10.33395/sinkron.v7i4.11783.
[18] IBM HR Analytics Employee. Available online: https://www.kaggle.com/pavansubhasht/ibm-hr-analyticsattrition-dataset