Harnessing Meta-Heuristics for Predictive Analytics: Enhancing Workplace Success through Advanced Algorithms

*

Srushti Sandigawad
Department of MCA
KLE Technological University
Hubballi 580031, Karnataka, India
srushtisandigawad@gmail.com

Ritul Mamdapur
Department of MCA
KLE Technological University
Hubballi 580031, Karnataka, India
ritulmamdapur17@gmail.com

Rashmi Benni
Department of MCA
KLE Technological University
Hubballi 580031, Karnataka, India
rashmi.benni16@gmail.com

Maheshwari M. Kittur
Department of MCA
KLE Technological University
Hubballi 580031, Karnataka, India
maheshwari.kittur@kletech.ac.in

***Abstract*—In today's dynamic professional world, predicting workplace success is crucial for fostering a productive and harmonious environment. The Big Five personality traits—Openness, Conscientiousness, Extra-version, Agreeableness, and Neuroticism are widely recognized for their impact on job performance and employee well-being. Openness involves creativity and change, while Conscientiousness reflects discipline and organization. Extra-version signifies sociability, Agreeableness involves compassion, and Neuroticism relates to negative emotions. The prediction of workplace success using the Big Five personality traits has become a prominent focus in the field of organizational psychology. Traditional predictive algorithms are unable to capture complex and multidimensional relationships inherent in personality data hence this study leverages biologically inspired meta-heuristic algorithms. Meta-heuristic algorithms are high-level problem solving algorithms which help finding optimal solutions. The classification algorithm Random Forest(RF) is used as predictive model for its robustness and ability to handle large datasets with multiple dimensions. This study conducts a comparative analysis of meta-heuristic algorithms, specifically Ant Colony Optimization(ACO) and Genetic Algorithm(GA), to enhance the predictive accuracy of workplace success models by finding optimal feature sets and weights. By providing a deeper understanding of how personality traits influence workplace success, this research contributes to countering toxic work environments and also helps improve team compositions. It offers valuable insights for organizations and human resource professionals focused on optimizing employee selection and development processes, ultimately fostering a more healthier and productive work environment.**

***Index Terms*—Workplace success prediction, personality, meta-heuristics, Genetic Algorithm (GA), Ant colony Optimization(ACO), Random forest(RF), optimal solutions, classification, nature-inspired, psychology.**

I. Introduction

In today's competitive corporate environment, predicting workplace performance has become essential to organizational development and employee management and recruitment. Traditional recruitment methods often overlook the significant influence of personality traits in job performance and success and place a heavy emphasis on cognitive ability and educational credentials [17]. The incorporation of personality assessment can be a major area of interest as it firms to improve productivity and better workplace environments [2] [13]. The five different personality traits provide a great framework for comprehending individual differences [16]. These characteristics offer insightful information on how employees may behave and perform in diverse work environments [9] [15]. Openness indicates the intellectual curiosity, creativity and is correlated with intelligence. Being conscientious is a great predictor of job performance since these people are typically industrious, dependable, and efficient. Extraversion encompasses sociability, zeal and assertiveness and often excel in roles demanding leadership, teamwork and communication. Agreeableness indicates qualities like cooperation, trust and selflessness. Neuroticism is used to characterize a person's susceptibility towards stressful emotions such as anxiety and moodiness. Individuals with high levels of neuroticism are undesirable at work as it has a negative impact [14], [18]. Recognizing the potential of these traits and using them in predicting workplace success, this study attempts to create reliable predictive models by utilizing sophisticated meta-heuristic algorithms. Meta-heuristic algorithms have a reputation for being effective in resolving optimization and combination problems, they present a viable path for improving predictive accuracy [5]. These algorithms are good at navigating large, complex datasets and find optimal solutions that may not be possible using conventional machine learning techniques [8], [22]. The two well-known meta-heuristic algorithms used are GA and ACO. Both ACO and GA are inspired by natural phenomena [10]. ACO is inspired by foraging behaviour of ants in search of food. shortest path. This behaviour is simulated in ACO in order to find optimal solutions to combinational problems. GA works on basis of the laws of natural selection. Over several generations it evolves a population of solutions through mechanisms like crossover, mutation and selection and works incredibly well in finding optimal or nearly optimal solutions. By applying GA and ACO to analyse the various five personality traits the goal of this research is to find the best patterns and correlations that can predict success in the workplace . The objective is to improve predictive models that enhance hiring processes and support organizational and employee development. The predictive model used will be Random Forest (RF) a powerful ensemble learning technique [6]. It uses number of decision trees while training and gives the mode of the classes as the final output [18], [21]. It is considered to be highly accurate, resistant to over fitting and able to manage large datasets with multiple dimensions. This study aims to provide a comparative analysis on the effectiveness between the two meta-heuristic algorithms by using Random Forest as a baseline [7]. It also aims to offer practical insights in the intersection of psychology and workplace success prediction. This topic is very important and relevant since it incorporates personality assessments to overcome the limitations of traditional recruitment methods which can result in more comprehensive and accurate predictions of employee performance and will ultimately improve the strategic management of human resources [19]. In the ever changing corporate world this study has the potential to transform hiring procedures, improve employee mental health and propel organizational success.

The paper is coherently organized. Section II talks about the initial background study. Section III describes the implemented work. Part IV presents the outcomes and remarks, Section V concludes the study, and Section VI discusses the study's potential scope.

II. RELATED WORK

Personality is the combination of cognitive abilities and behaviours that are acquired from various biological and environmental factors. The MBTI(Myers Briggs Type Indicator) model, Big Five model and Theory of personality types are widely recognized frameworks and offer a structured approach to classify various personality types [11], [13]. The study of personality is a significant area within psychology and a important factor in evaluating candidates for corporate jobs. For job seekers, knowing your strengths helps you to improvise and showcase your talents to employers [20]. Process of recruitment has evolved from traditional methods of recruitment to more advanced levels based on the candidate's intellectual and cognitive abilities. But these methods do not consider the candidate's personality and behaviour as a factor. Hence including them can assist recruiters to shortlist candidates better equipped to the job requirements [19]. Numerous studies have explored the integration of personality assessments into recruitment using machine learning (ML) techniques. Kavya P utilized Analysis of Variance(ANOVA) to analyze differences among groups, K-means clustering to group students, and Support Vector Machine(SVM) for classification of student personalities. It highlights the utility of combining statistical methods with clustering and classification algorithms to provide a comprehensive analysis of personality data [12]. Likewise, Atharva Pansare's study leveraged multiple ML techniques, including Logistic Regression, Random Forest, SVM, and XGBoost, for personality prediction based on MBTI personality assessment tool. The study's results indicate that advanced ensemble methods like XGBoost can capture sophisticated patterns in personality traits, making it highly suitable for predictive tasks [13]. Expanding on this, the study by Didi Supriyadi also compared various ML algorithms for student personality classification. This comparison demonstrates the versatility and robustness of these techniques in processing and analysing personality data [18]. The findings from Pansare and Supriyadi's research complement to one another as they both emphasize the effectiveness of ensemble methods and advanced algorithms in personality prediction Supriyadi's findings supported Random Forest and SVM which are particularly known for their ability to handle high-dimensional data and their robustness against over-fitting [13], [18]. Conversely, Hima Vijay used K-means clustering to classify different personalities based on Big five model. The unsupervised learning technique K-means clustering effectively categorizes data points based on similarities and differences [20]. Additionally, another paper stresses the application of K-means clustering to improve system productivity and efficiency while also adding Principal Component Analysis(PCA) for visualizing cluster predictions which helps to minimize the dimensions of the data, improving interpretability of complex data [8]. In Md. Thoufiq Zumma's research, pre-processed MBTI dataset containing the four-letter MBTI personality descriptions was compared with the Big Five model. This comparison established links between the dimensions of both models by providing insights into how different personality frameworks relate to each other. Comparative studies like these are important for blending insights from multiple models and improving the precision of personality assessments [22]. Rohit HV highlighted the potential of social media data in personality prediction using Facebook profiles as the primary source. Random Forest algorithms which are known for their accuracy and interpretability were leveraged to analyse online social fingerprints [14]. Overall, these studies illustrate the diverse applications of various ML models in personality prediction. Traditional models have been extensively trained and evaluated, showcasing their strengths and limitations. However, this research aims to integrate the Big Five personality model, a widely recognised and validated model for personality assessment with meta-heuristic algorithms to explore a large search space and find near-optimal solutions, also solving combinatorial optimization problems making it suitable for the complex task of personality prediction. This combination will also offer deeper insights into the underlying personality traits, contributing to more personalized and effective applications in various domains such as human resources, mental health, and personalized marketing.

III. PROPOSED WORK

The proposed research study investigates comprehensive meta-heuristic algorithms ACO and GA for enhancing the predictive modeling of workplace success, in conjunction with Random Forest as the predictive model. The principal goal is to apply and compare these methodologies in a systematic manner with the goal of improving the forecast accuracy of workplace success models. It seeks to leverage advanced optimization techniques to ensure high quality solutions and increase the reliability and applicability of the predictive model for workplace success.

A. System Architecture

The system architecture illustrates the application of meta-heuristic algorithms, specifically GA and ACO, to optimise predictive modelling of workplace success. The input data crucial to this predictive modeling endeavor consists of meticulously generated synthetic data. This synthetic data encapsulates the comprehensive spectrum of personality test scores covering the Big Five traits and a Performance score of the employee based on these traits. The input data is first preprocessed in order to handle missing values. Subsequently, the algorithms are employed and their hyper parameters are fine-tuned through optimisation techniques to enhance the fitness function. This function evaluates the solutions on the basis of accuracy of the Random Forest classifier. When the solution reaches the necessary accuracy, the predictive model is evaluated and trained using different evaluation metrics. This approach yields informative data when compared to alternative optimisation strategies to find the most accurate and feasible solution.

B. Genetic Algorithm

GA is a search based meta-heuristic algorithm which works on the basis of natural selection [4]. It is frequently used to produce optimal solutions to optimization and combinatorial problems by depending on its bio inspired operations like mutation, crossover and selection [3]. It belongs to the wider family of evolutionary algorithms. GA is used to optimize the weights assigned to the big five traits to maximize the accuracy of the Random Forest predictive model. A population of weights is generated where each individual denotes a set of weights and represented as a genome. A fitness function evaluates these individuals by using the sum of weights, normalizing the scores and accuracy of the Random Forest Classifier. The best individuals are selected and mutated to produce new offspring individuals. Recombination is used to combine pairs of parents and find most optimal solutions. The population is then evolved over multiple generations and the best individual of the final generation is identified . This best set of weights is used to train the Random forest classifier to determine its performance.

- **Initialization**
  Generate a population of N individuals, each represented by:
  $$\text{Individual}_i = [w_1, w_2, ..., w_n]$$
  where n is the number of genes.
- **Selection**
  Select individuals i using tournament j and selection of size k:
  $$\text{P(selected)(Individual}_i) = \frac{1}{k}\sum_{j=1}^{k} \delta(\text{i}, \text{j})$$
- **Crossover** Use two-point crossover with parents P_1 and P_2, and crossover points cp_1, cp_2:
  $$\text{Offspring}_1 = [\text{P}_1[1 : cp_1], \text{P}_2[cp_1 : cp_2], \text{P}_1[cp_2 : end]]$$
  $$\text{Offspring}_2 = [\text{P}_2[1 : cp_1], \text{P}_1[cp_1 : cp_2], \text{P}_2[cp_2 : end]]$$

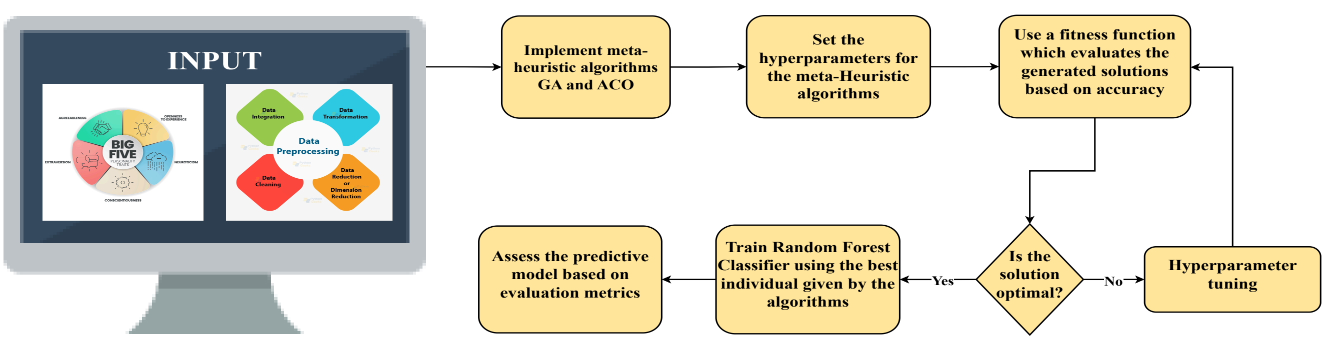

Fig. 1: System Architecture

- **Mutation**
  Introduce random changes with probability p_m:

$$\text{Mutated Gene} = \begin{cases} \text{random\_uniform}(-1, 1) \\ \text{with probability } p_m \\ \text{original gene} \\ \text{with probability } 1 - p_m \end{cases}$$

- **Fitness Function**
  Calculate the fitness of an individual based on its performance score, normalized score, and target variable:

$$f(\text{Individual}_i) = \text{accuracy}$$

where:

$$\text{Performance Score}_j = \begin{aligned} & w_1 \cdot x_j^{(1)} + w_2 \cdot x_j^{(2)} + w_3 \cdot x_j^{(3)} \\ & + w_4 \cdot x_j^{(4)} + w_5 \cdot x_j^{(5)} \end{aligned}$$

$$\text{Normalized Score}_j = 100 \times \frac{\text{score}_j - \min(\text{score})}{\max(\text{score}) - \min(\text{score})}$$

$$y_{\text{train\_pred}} = \begin{cases} 0 & \text{if Normalized Score}_j \leq 33 \\ 1 & \text{if } 33 < \text{Normalized Score}_j \leq 66 \\ 2 & \text{if Normalized Score}_j > 66 \end{cases}$$

- **Termination**
  The GA algorithm ends when a predefined criteria such as set number of generations is met.

C. Ant Colony Optimization Algorithm

Ant Colony Optimization algorithm is another algorithm which works on the basis of the foraging behaviour of ants and their natural ability to search for the shortest path between their nest and the food source [1]. As the name suggests it is used for solving optimization problems especially in combinatorial optimization domain. It imitates the pheromone secretion and evaporation by ants. ACO is used to optimize the weights assigned to the traits to maximize the accuracy of the Random forest classifier. Artificial ants which represent the weights move through the solution space. A pheromone matrix is updated based on the fitness of the ants. The ant movement is repeated for a fixed number of iterations and the pheromone at the sub optimal solution paths is evaporated. The best weight combination is then extracted based on the concentration of pheromone representing the optimal solution for the Random Forest Classifier.

- **Initialization**
  Initialize the ACO algorithm with the following parameters:
  - n_{ants}: Ant colony size.
  - $n_{\text{iterations}}$: Number of iterations.
  - ρ: Pheromone decay rate.
  - α:Pheromone importance.
  - β: Heuristic information importance.
  - τ_{ij}: Pheromone level on the path (i, j), initialized to a constant value.
- **Ant Movement**
  Ants move based on the probability of choosing the next node:

$$p_{ij} = \frac{\tau_{ij}^{\alpha} \eta_{ij}^{\beta}}{\sum_{k \in \text{allowed}} \tau_{ik}^{\alpha} \eta_{ik}^{\beta}}$$

  where η is the heuristic information.
- **Fitness Evaluation**
  Evaluate the fitness of each ant's solution based on its ability to predict success using a fitness function.

$$f(\text{Ant Solution}) = \text{Accuracy}$$

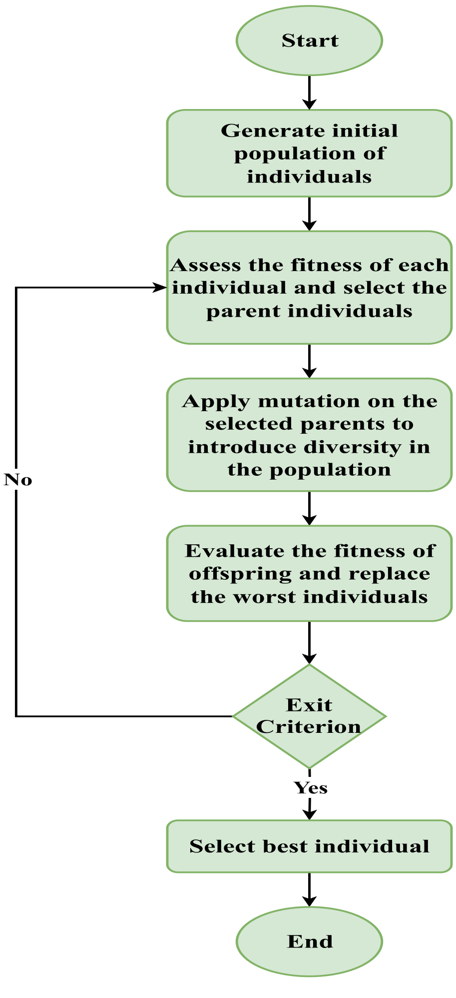

Fig. 2: Genetic Algorithm

- **Pheromone Deposit**
  After finding a solution, each ant deposits pheromone on path based on the fitness of its solution:
  $$\tau_{ij} = \tau_{ij} + \frac{Q}{f(\text{Ant Solution})}$$
  where Q represents the intensity of deposited pheromone.
- **Pheromone Update**
  Update pheromone levels on edges using:
  $$\tau_{ij} = (1-\rho)\tau_{ij} + \frac{1}{\text{max\_fitness}}$$
  where ρ is the rate of decaying the pheromone and max_fitness is the maximum fitness value found in the current iteration.
- **Iteration Convergence**
  Repeat ant movement, fitness evaluation, pheromone deposit, and pheromone update for a fixed number of iterations.

D. Random Forest Classifier

Random Forest is a form of ensemble learning technique which builds multiple decision trees each of which are trained on distinct set of data. It is based on the idea that multiple models combined produce a more robust and accurate prediction. It gives the class that is the most repeated prediction of all the decision tress during classification and mean of the prediction of individual trees for regression. The ensemble classifier is used to classify the workplace success into three categories: 'Low', 'Medium', and 'High'. It is trained on a new target created from the normalized scores based on the weights assigned by the meta heuristic algorithms. The algorithms ACO and GA produce random set of weights. The algorithms then find the best set of weights using a fitness function. The best combination of weights is then used to train the classifier on the testing dataset.

- **Bootstrap Sampling**
  Create multiple smaller datasets from the original training set:
  $$D_i = \{(x_1, y_1), (x_2, y_2), \ldots, (x_{N_i}, y_{N_i})\}$$
  in which D_i is the i-th dataset sample, (x_j, y_j) are the data points, and N_i is the number of samples in D_i.
- **Building Decision Trees**
  For each dataset sample, build a decision tree:
  $$T_i = \text{build\_tree}(D_i)$$
  where build_tree is the function that constructs a decision tree from the sample D_i.
- **Random Feature Selection**
  During the construction of each tree a small set of features is chosen at each node:
  $$F_i = \text{random\_subset}(F, k)$$
  where F is the set of all features and k is the number of features.

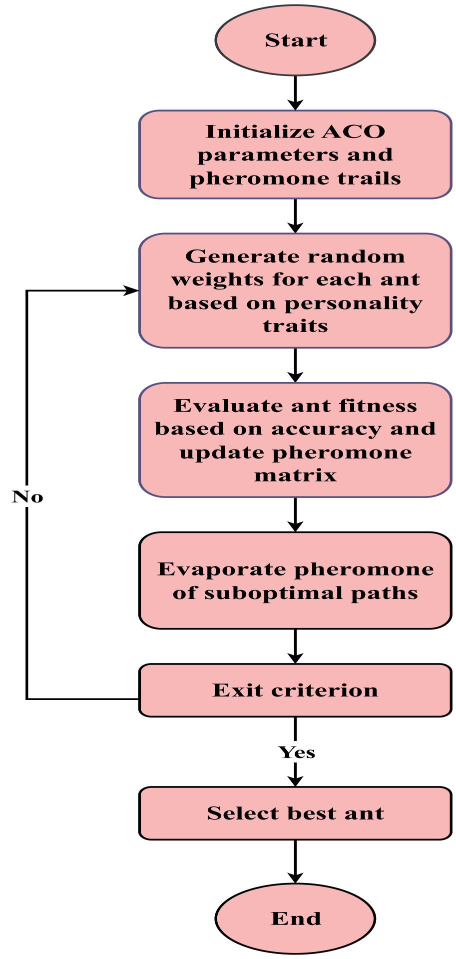

Fig. 3: The Ant Colony Optimization Algorithm

- **Splitting Criterion**
  Use the most ideal feature to divide the node based on Gini impurity or entropy:
  $$\text{Gini impurity:} \quad G(t) = 1 - \sum_{i=1}^{C} p_i^2$$
  $$\text{Entropy:} \quad H(t) = -\sum_{i=1}^{C} p_i \log(p_i)$$
  where C is the total number of classes,$G(t)$ is the Gini impurity of node t, $H(t)$ is the entropy of node t, and p_i is the likelihood of an element being classified i in node t. The Gini impurity measures the probability of incorrect classification, while entropy measures the disorder of the node.

- **Aggregation of Trees**
  The class predicted is the mode of the predictions from all the individual trees:

  For classification:
  $$\hat{y} = \text{Mode}(\{T_i(x)\}_{i=1}^{M})$$

  where $\hat{y}$ is the output class label, and M is the total count of trees.

IV. RESULTS AND DISCUSSION

This work assesses how well ACO and GA improved Random Forest model for predicting workplace success in three different test scenarios. The amount of the dataset and the setups of the hyperparameters differed for each test scenario. A dataset including 1000 samples was used in the first test case; the dataset was enlarged to 2000 in the second test case; and GA parameters like number of generations,crossover rate and mutation rate while in ACO number of iterations,pheromone evaporation rate were adjusted while the dataset size remained at 1000 in the third test case to evaluate the trade off between hyperparameters and performance of the algorithms.

EXPERIMENT I		
Algorithms	GA	ACO
Dataset Size	1000	1000
Initial Fitness	0.81	0.70
Average Fitness	0.885	0.718
Best Fitness	0.9067	0.6600
Number of Iterations	20	20
Accuracy	0.91	0.66
Precision	0.91	0.70
Recall	0.91	0.66
F1 Score	0.91	0.67

TABLE I: Comparison of GA and ACO in Experiment I

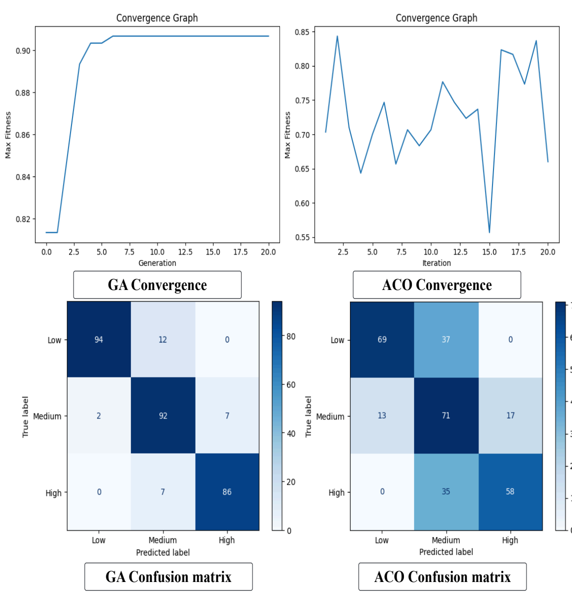

Fig. 4: EXPERIMENT I

In the experimental setup I, GA demonstrated an impressive increase in fitness values in the first test instance, from an initial fitness of 0.8133 to a best fitness of 0.9067. With an accuracy of 0.91 and balanced evaluation parameters across different classes, the classification performance was as outstanding. ACO, on the other hand had an initial fitness of 0.7033 was unable to sustain a steady increase in fitness values; as a result, it ultimately achieved a final fitness of 0.66 .The best individual identified by GA had weights [0.3457, 0.9585, 0.4595, 0.4709, -0.5186], reflecting a well-optimized set of parameters contributing to the high fitness value of 0.9067 meanwhile ACO's best individual had weights [-0.1512, 0.7207, 0.0843, 0.3793, -0.0614] and an accuracy value of 0.66. GA's classification metrics were also superior for all the class labels . The top graphs show the convergence behavior of both GA and ACO.GA displays a constant and quick gain in fitness that stabilizes at a high level, whereas ACO has more fluctuation and less stability in fitness levels. The confusion matrices indicate that ACO displays more errors and less accurate predictions than GA,especially in the "Medium" and "High" categories.

EXPERIMENT II		
Metrics	GA	ACO
Dataset Size	2000	2000
Initial Fitness	0.745	0.878
Average Fitness	0.88375	0.75133
Best Fitness	0.92	0.58
Number of Iterations	20	20
Accuracy	0.92	0.58
Precision	0.92	0.68
Recall	0.92	0.58
F1 Score	0.92	0.58

TABLE II: Comparison of GA and ACO in Experiment II

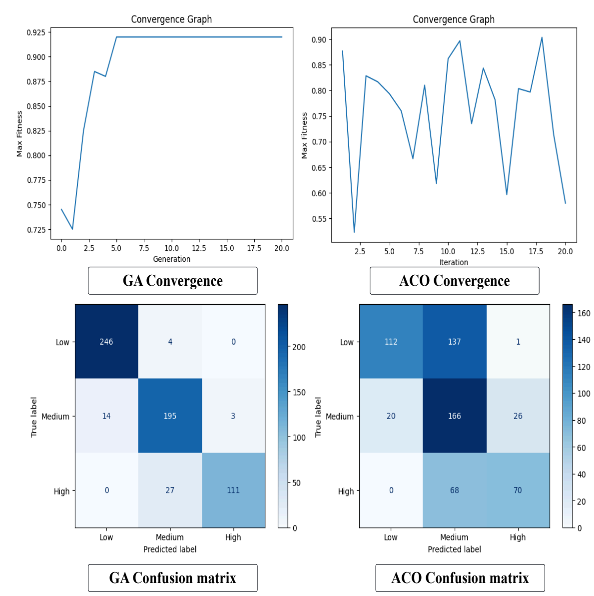

Fig. 5:
EXPERIMENT II

In the second experimental setup even with a larger dataset, the trend persisted as GA significantly outperformed ACO in fitness values. In this test case, the best individuals weights were [0.8805, 0.8887, 0.9071, 0.6792, -0.5613], showcasing GA's ability to find optimal solutions even with an increased dataset size and ACO's the weights were [0.2815, 0.6268, -0.3011, 0.1906, -0.6587]. The classification performance with GA was notable, with an accuracy of 0.92 and balanced precision, recall, and f1-scores across different classes. In stark contrast, ACO struggled to maintain a consistent improvement in fitness values,it had an initial fitness of 0.87 but could not maintain it and had a final fitness of 0.58. As indicated in the convergence graphs, GA had an initial fitness of 0.74 and it was stabilized to 0.92 in the later generations. ACO had a higher initial fitness of 0.87 but could not stabilize it in the later iterations. GA shows high accuracy in classifying the 'low' and 'medium' class labels with some miss classifications in the 'high' class label. The experiment findings demonstrated the robustness and scalability of the GA due to its stable convergence nature for larger datasets while ACO suffers premature convergence.

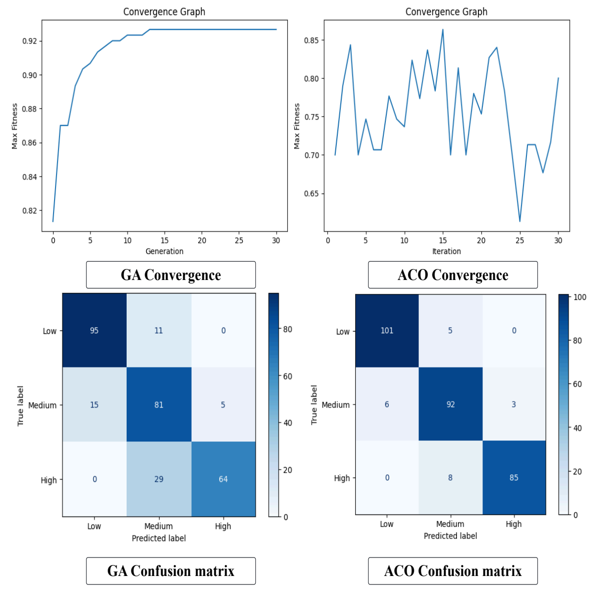

Fig. 6: EXPERIMENT III

EXPERIMENT III		
Metrics	GA	ACO
Dataset Size	1000	1000
Initial Fitness	0.813	0.70
Average Fitness	0.908444	0.751667
Best Fitness	0.9267	0.8
Number of Iterations	30	30
Accuracy	0.93	0.80
Precision	0.93	0.82
Recall	0.93	0.80
F1 Score	0.93	0.80

TABLE III: Comparison of GA and ACO in Experiment III

The final experimental setup III demonstrated the robustness of GA even more by adjusting hyperparameters. With a best fitness of 0.9267, GA proved that it could successfully use parameter changes to improve model performance.The number of generations were increased to 30 and crossover rate and mutation rate were set to 0.6 and 0.3 respectively while in ACO the number of ants were increased to 50 and iterations were set to 30 with decay rate 0.2. Compared to the best fitness of 0.80 for ACO, the classification accuracy of 0.93 for GA was significantly higher than ACO. Whereas ACO's performance was more unpredictable and less reliable with the best weights of [0.419 , 0.354, 0.174, 0.441, -0.156]. GA showed steady and consistent increases in fitness values across generations in all test situations. GA's adaptability with the best individual having weights [0.3523, 0.8155, 0.5075, 0.4709, -0.3272], which led to the highest fitness value. The convergence behaviour of both the algorithms in 30 generations or iterations remained the same.

There are a few important variables that contribute to GA's exceptional performance. First of all, GA can efficiently explore and exploit the solution space thanks to its global search capabilities, which is made possible by its iterative process of selection, crossover, and mutation. Through this procedure, GA is able to steer clear of local optima and gradually increase fitness values, which results in more ideal solutions. Furthermore, as shown by its performance in the second test scenario, GA can scale well due to its adaptation to larger datasets. The third test case's hyperparameter adjustments provided more evidence of GA's adaptability and durability in terms of model parameter optimization.On the other hand, ACO struggled to find stable solutions and was prone to exhibiting fluctuations in fitness values, which negatively impacted its performance. ACO performs well in some optimization scenarios, but not always; this proved especially true when using larger datasets and adjusting hyperparameters. This inconsistency can be ascribed to ACO's reliance on pheromone-based pathways, which may not be as effective in negotiating the complicated solution space of Random Forest optimization for workplace success prediction.Furthermore, combinatorial problems—which entail determining the best combination of variables that produce the optimal solution—are especially well-suited for GA. This benefit stems from GA's population-based methodology, which preserves a wide range of viable answers and improves its capacity to comb over a large search area. By introducing fresh and varied parameter combinations, the crossover and mutation operations enable efficient navigation of the intricate world of combinatorial optimization problems. Overall, GA's consistent improvements, scalability, and stability make it a more suitable choice for enhancing Random Forest models in this application compared to ACO. GA is a useful tool for predicting workplace success because of its strong global search capabilities and effective exploration processes, which allow it to acquire greater fitness values and superior classification performance.The best individuals found by GA also show how well it can solve combinatorial optimization problems by identifying efficient parameter combinations that greatly improve model performance.

V. Conclusion

The article findings and the comparative experimental setup shows that GA consistently outperforms ACO in terms of fitness, and better improves the accuracy of the predictive model.GA's evolutionary mechanisms of selection, crossover, and mutation—which effectively explore and utilize the search space to identify optimal solutions—are responsible for its excellent performance.It is particularly effective for combinatorial problems due to its ability to maintain genetic diversity preventing premature convergence and explores a vast solution space effectively. On the other hand, ACO's performance, is limited by its dependence on pheromone trails, which may result in less-than-ideal outcomes if they are not counterbalanced with sufficient exploration. Thus, as the experimental results show, GA is a more appropriate option for complicated optimization issues due to its adaptable and versatile character.

VI. Future work

The future objectives of this study would be to assess the robustness of the proposed work, examine the efficacy of additional meta-heuristic algorithms and their hybridisation, to broaden its scope to include more complex scenarios such as changing job roles, dynamic team compositions, and diverse workplace environments. Evaluate the scalability of the proposed approach by applying it to real-world applications as employee retention strategies, talent acquisition, and productivity enhancement. The combination of machine learning methods and meta-heuristic algorithms could be investigated further in order to improve prediction accuracy and adapt to the quickly evolving dynamics of contemporary workplaces. Furthermore, assessing these algorithms' effects in various organizational contexts and industries may offer important new perspectives on their applicability and generalizability.

References

[1] J. E. Smith A. E. Eiben. An article on introduction to evolutionary computing. *Springer Berlin, Heidelberg natural computing series*, 2013.

[2] Raj Agarwal, Roshni Mukherjee, Nitin More, and Suvarna Pawar. Big five (o.c.e.a.n): A new way to recruit. In *2023 1st International Conference on Cognitive Computing and Engineering Education (ICCCEE)*, pages 1–7, 2023.

[3] Seán McGarraghy Anthony Brabazon, Michael O'Neill. An article on natural computing algorithms. *Springer Berlin, Heidelberg natural computing series*, 2015.

[4] Rashmi Benni. A comparative study of evolutionary and swarm intelligence algorithms for job scheduling on identical parallel machines. *Tuijin Jishu/Journal of Propulsion Technology*, 44:3734–3745, 10 2023.

[5] Rashmi Benni, Shashikumar Totad, Karibasappa K G, and Sachin Karadgi. Search space optimization for autonomous mobile robots using meta-heuristics. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–8, 2023.

[6] D.N.A. Dissanayake, K.S. Ranasinghe, and R.A.H.M. Rupasingha. Forecasting employee personality changes in the post covid-19 era: Ensemble learning approach. In *2023 5th International Conference on Advancements in Computing (ICAC)*, pages 822–827, 2023.

[7] Bandi Ganesh and Dr. K. Anitha. Implementation of personality detection and accuracy prediction for identification of fake and true news using decision tree and random forest algorithms. In *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, pages 1–5, 2022.

[8] Mansi Goel, Sanjeev Pippal, Sansar Chauhan, and Rajiv Kumar. Personality prediction and analysis using clustering. In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, pages 271–275, 2021.

[9] Mariam Hassanein, Sherine Rady, Wedad Hussein, and Tarek F. Gharib. Predicting the big five for social network users using their personality characteristics. In *2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS)*, pages 160–164, 2021.

[10] Md. Saddam Hossain Mukta, Jubaer Ahmad, Akib Zaman, and Salekul Islam. Attention and meta-heuristic based general self-efficacy prediction model from multimodal social media dataset. *IEEE Access*, 12:36853–36873, 2024.

[11] Zuriani Mustaffa, Nur Alia Shahira Mohd Zaidi, Ferda Ernawan, Haithm Elhadi, and Muhammad Malik Hakim. Personality predictive analysis based on artificial neural network. In *2022 6th International Conference on Informatics and Computational Sciences (ICICoS)*, pages 105–110, 2022.

[12] Kavya P and Kanchana V. Student personality analysis in blended mode using big five. In *2023 International Conference on Network, Multimedia and Information Technology (NMITCON)*, pages 1–5, 2023.

[13] Atharva Pansare, Prabhat Panwar, and Pranali Kosamkar. Personality prediction with natural language processing using questionnaire responses. In *2022 IEEE Pune Section International Conference (PuneCon)*, pages 1–6, 2022.

[14] G. V. Rohit, K Rakesh Bharadwaj, R Hemanth, Bariti Pruthvi, and M.V. Manoj Kumar. Machine intelligence based personality prediction using social profile data. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pages 1003–1008, 2020.

[15] Xinyu Shui, Yiling Chen, Xin Hu, Fei Wang, and Dan Zhang. Personality in daily life: Multi-situational physiological signals reflect big-five personality traits. *IEEE Journal of Biomedical and Health Informatics*, 27(6):2853–2863, 2023.

[16] Brijesh Kumar Singh, Mansi Katiyar, Shefali Gupta, and Nikam Gitanjali Ganpatrao. A survey on: Personality prediction from multimedia through machine learning. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pages 1687–1694, 2021.

[17] G. Sudha, Sasipriya K K, Sri Janani S, Nivethitha D, Saranya S, and Karthick Thyagesh G. Personality prediction through cv analysis using machine learning algorithms for automated e-recruitment process. In *2021 4th International Conference on Computing and Communications Technologies (ICCCT)*, pages 617–622, 2021.

[18] Didi Supriyadi, Purwanto, and Budi Warsito. Performance comparison of machine learning algorithms for student personality classification. In *2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, pages 73–78, 2022.

[19] Lakshyajit Thapa, Aniket Pandey, Deepanshu Gupta, Ayush Deep, and Rakesh Garg. A framework for personality prediction for e-recruitment using machine learning algorithms. In *2024 14th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, pages 1–5, 2024.

[20] Hima Vijay and Neenu Sebastian. Personality prediction using machine learning. In *2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS)*, pages 1–6, 2022.

[21] Yuchen Wang. Personality type prediction using decision tree, gbdt, and cat boost. In *2022 International Conference on Big Data, Information and Computer Network (BDICN)*, pages 552–558, 2022.

[22] Md. Thoufiq Zumma, Jerin Akther Munia, Dipankar Halder, and Md. Sadekur Rahman. Personality prediction from twitter dataset using machine learning. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–5, 2022.