mLAC Journal for Arts, Commerce and Sciences (m-JACS)
Volume 4, No.5, June 2026, P 1-6
ISSN: 2584-1920 (Online)

PREDICTION OF CANDIDATE JOINING AND RECRUITING PROCESS USING MACHINE LEARNING

Mohitha M[1], Meghana B S[1*]
[1]Dept. of MCA, P.E.S College of Engineering, Mandya, India1.
Corresponding author email address: meghana@pesce.ac.in

Paper Received: 23.01.2026 | Revised: 28.04.2026 | Accepted: 29.05.2026

DOI: https://doi.org/10.59415/mjacs.361

Abstract

In the era of advanced technology and digitalization, the use of machine learning makes it possible to increase the efficiency of the recruitment process. This project work aims to identify various variables that directly play a role in recruitment. The work also addresses the HR sphere problems and proposes a solution related to the use of machine learning. The proposed methodology includes an analysis of the characteristics which are dominated by past candidates. The data obtained suggests that machine learning can bring significant benefits to the recruitment process, such as reducing the cost and time for selection, reducing emotional factors, and increasing accuracy. Overall, this proposed work provides valuable insight into the potential of machine learning in recruitment and highlights how it can be used to facilitate the recruitment process. Employee recruitment process is one the complex tasks in the current generation. There are many reasons where candidate may reject the offer given by the company. So it is very difficult to identify which candidate may join and which candidate may reject the offer. In this proposed system we build an automation for prediction whether candidate joins to the company or not. HR datasets collected from the reputed websites like www.kaggle.com, www.dataworld.com etc... In this proposed system we develop automation for companies to know the candidates joining. Company joining prediction is done using Machine Learning algorithms. Efficient ML algorithm used to process the HR data. Proposed system is a browser-based application meant for any organization and developed using Microsoft technologies such as Visual Studio, C# and SQL Server.

Keywords: Recruitment, Candidate Joining, Machine Learning, Supervised Learning, Human Resources, Real-Time Prediction.

1. INTRODUCTION

Recruitment process in any organization is a complex process which requires more time, human resources and more expensive. Initially every organization gives recruitment advertisements or ads on job openings through different platforms such as social media, TV, newspaper and business websites. In any organizations the process is long and eligible candidates will apply for the jobs and initially HR team will manually sit and shortlist the suitable candidates for the interview process and interview dates will be fixed. Interview process will be based on the type of the job, many rounds of interviews will be there like in IT company interview rounds includes aptitude, technical round, group discussion, HR round etc.… candidates should pass through all interview rounds for the selection. After the interview process list of suitable candidates will be decided and selected by the HR team based on the job counts and designations. All these selected candidates will be intimated regarding the same and offer letters will be sent to the selected candidates. Here there is an issue, based on the company
requirement suitable candidates will be selected, but all selected candidates may not join to the company due to one or the other reason. So it important to identify such candidates for the company sake or else again company should start the new recruitment process to fulfill their recruitment policy again which requires more time and money.
Machine learning techniques applied in all fields to process the existing data and to train the system so as get the accurate results. In machine learning there are 3 different types, supervised learning, unsupervised learning and semi

.
supervised learning techniques. In this project work to predict the candidate joining we use efficient machine learning algorithms. Supervised learning algorithms such as Bayesian algorithm, ID3 algorithm, Random forest algorithm or decision tree algorithm used to process the training datasets downloaded from kaggle.com and predicts candidate joining. Many existing research works on this topic build machine learning models using python, r language, weka tool etc... But these works are not real time implementations and not suitable for real time. All research works used static datasets and shown results, many papers got around 80% results and few papers got around 90% results. In this project we build machine learning model with real time application useful for the organization in recruitment process to identify candidate joining. If company knows who are all selected candidates will join job, company can properly plan for further process and can reduce the investments in recruitment process. Proposed system predicts the candidate joining in advance prior to candidate joining in real time. Efficient Microsoft technologies used for the project development such as visual studio and SQL server.

2. RELATED WORK

1.The paper titled "An Intelligent Career Guidance System using Machine Learning" (2021) by Vignesh S., Shivani Priyanka C., Shree Manju H., and Mythili K. introduces a computer-based career counseling tool aimed at assisting students in selecting the appropriate academic stream after completing their higher secondary education. The system relies on K-Nearest Neighbors (KNN) to evaluate skills based on objective test results and match them with the most compatible engineering departments. It identifies students' core strengths and suggests suitable fields like Computer Science, Mechanical, or Electrical Engineering. While the system effectively guides students in academic planning, it does not consider employment behavior or recruitment outcomes, and hence, does not contribute toward predicting candidate-joining.

2. In the study "Predicting Students' Employability using Machine Learning Approach" (2020), authors Cherry D. Casuat and Enrique D. Festijo propose a machine learning model that estimates the employability of students based on academic and internship performance. Algorithms such as Decision Trees, Random Forest, and Support Vector Machines were tested on datasets containing students' GPA, job interview scores, and on-the-job training ratings. The system demonstrates how data-driven approaches can forecast employability prior to graduation. However, it does not extend to real-time recruitment or joining prediction, and its implementation remains largely academic without integration into industry workflows.

3. The paper "Application of Data Mining in Predicting College Graduates Employment" (2021) by Shouwu H., Xiaoying Li, and Jia Chen focuses on applying CART decision tree and Random Forest was applied to forecast job placement outcomes for college graduates in China. Using factors like academic performance and graduation status, the model achieved a prediction accuracy of 81%. While the methodology is effective in analyzing graduate outcomes and identifying patterns in employment, it is primarily statistical in nature and lacks the predictive scope needed to assess post-offer candidate behavior or joining likelihood.

4. Akku George Saju, in his 2018 paper titled "Future Job Prediction in Trivandrum Using Machine Learning Techniques", presents an interactive web application that forecasts job availability and salary trends across various sectors like IT, finance, and administration. By applying machine learning algorithms including Decision Trees, Linear Regression, and Neural Networks, the system helps job seekers plan their career path and understand required qualifications. Although useful for students and job aspirants, the work is focused on future job projections rather than recruitment outcome prediction, and does not assess candidate Joining behavior after job offers.

5. The paper "Appropriate Job Selection Using Machine Learning Techniques" (2023) by Md. Ashikur Rahman Khan et al. presents a classification model to determine job suitability based on an individual's experience, preferences, and skillset. Using algorithms like Decision Tree, Gaussian Naïve Bayes, Random Forest, and Support Vector Machine, the study achieved high accuracy in recommending appropriate job roles. While the approach supports fresh graduates in selecting ideal positions, it does not forecast whether a candidate will join a company post-offer, nor does it integrate business rules or real-time deployment features necessary for modern HR systems.

6. The study "Online Recruitment System and Personality Prediction" (2021) by Aditi Chandrashekar et al. proposes a recruitment framework that integrates aptitude testing with personality analysis based on the Big Five Traits model. Logistic regression is used to classify candidates into behavioral categories to support decision-making in hiring. Although this system brings personality insights into recruitment, it does not address offer acceptance prediction or time-sensitive decision-making needed in corporate hiring scenarios.

7. In "Predictive Model on Employability of Applicants and Job-Hopping using Machine Learning" (2022), the authors introduce a data-driven approach to forecast employability and job-hopping tendencies using classifiers like Naive Bayes, Decision Tree, and SVM. While the model adds value by identifying candidates likely to switch jobs frequently, it mainly focuses on long-term job behavior rather than predicting immediate joining outcomes after offer

.
rollout.
8. In "Enhancing Job Recruitment Prediction through Supervised Learning and Structured Intelligent Systems" (2022), the authors develop a machine learning model using a government dataset from Akwa Ibom State to predict applicant employability. Decision Tree and Random Forest are among the top-performing algorithms. 2.The findings show the strength of tree-based models in streamlining large-scale government recruitment. 2.However, the model does not provide real-time insights on whether a selected candidate is likely to join, and its design may not directly apply to private sector recruitment systems.
9. The research article "Analysis of the Effectiveness of Using Machine Learning Algorithms to Make Hiring Decisions" (2023) by Smelyakov et al. evaluates decision tree-based models, including Random Forest, on a dataset of 614 job applicants. Performance is measured using precision, recall, and F1-score. While the study highlights ML's role in improving candidate classification, it does not account for offer-related decisions or real-time hiring outcomes.

3. METHODOLOGY

EXISTING SYSTEM

The current recruitment process in most companies is manual, involving resume screening, interviews, and offer rollouts without any predictive insight into whether candidates will actually join. HR teams rely solely on experience and assumptions, making the process time-consuming, costly, and often inefficient. If selected candidates decline offers, companies must restart the hiring cycle, leading to delays and wasted resources. There is no automated system to analyze candidate behavior or predict joining likelihood, resulting in increased workload and reduced hiring accuracy.

Disadvantages:

- Entire recruitment process is time-consuming and resource-intensive.
- No predictive mechanism to identify candidates likely to reject offers.
- Requires repeated cycles of hiring if candidates back out.
- Increases cost and administrative workload.
- Prone to human bias and subjective decision-making.
- Lacks integration with intelligent tools for data analysis and prediction.
- No real-time or dynamic support to aid HR in planning or forecasting candidate behaviour.

4. PROPOSED WORK

The proposed system aims to automate and predict whether a selected job candidate will join the company or not, using a real-time software application. This system is designed as a full-fledged GUI-based solution accessible through common web browsers like Chrome, Opera, and Edge. It processes various candidate parameters such as gender, age, previous experience, job type, designation, offered salary, and location to make accurate predictions. The system leverages supervised machine learning algorithms to analyze historical recruitment data and forecast candidate joining behavior. It is developed using robust Microsoft technologies like Visual Studio and SQL Server, ensuring smooth handling of real-time and dynamic datasets. This intelligent tool is intended to support HR teams by reducing uncertainty and improving recruitment efficiency.

Advantages:

- Automates candidate joining prediction, reducing manual effort.
- Provides real-time decision support for HR teams.
- Enhances accuracy using ML algorithms trained on historical data.
- Handles large, dynamic datasets efficiently.
- Minimizes recruitment delays and cost by forecasting dropouts in advance.

.

- Accessible via standard web browsers, making it user-friendly and scalable.
- Built using reliable and industry-proven tools like Visual Studio and SQL Server.

5. IMPLEMENTATION

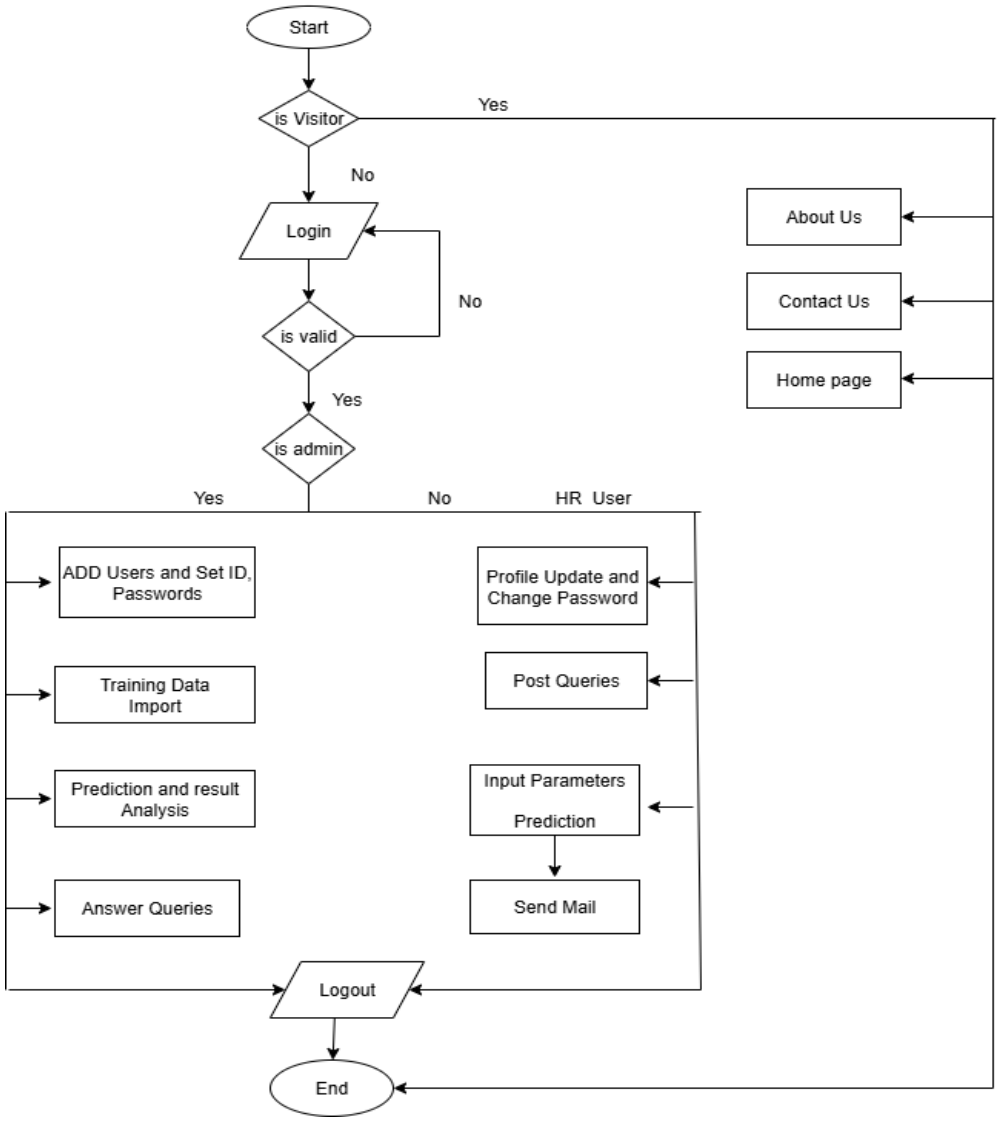
5.1 Operation of Random Forest

The Random Forest (RF) algorithm is used in this project to predict whether a selected candidate will join the organization after receiving a job offer. The steps involved in the algorithm's operation are as follows:

1. Random Sampling and Attribute Selection:A random seed is initialized to sample a subset of candidate records from the training dataset. This ensures class balance is preserved. From this sample, a random subset of candidate attributes (e.g., experience, notice period, age) is selected to avoid computational overhead and reduce overfitting.
2. Attribute Limitation per Tree:If the dataset contains M attributes, only R (where R < M) attributes are randomly chosen to construct each decision tree. This maintains tree diversity within the ensemble model.
3. Decision Tree Construction:Using the selected attributes, each tree is built by recursively identifying the best split at each node using the Gini index. This process continues until a stopping condition is reached—typically when leaf nodes become too small to further divide.
4. Majority Voting:Each tree in the forest outputs a classification—either "Joined" or "Not Joined". The final prediction is determined by majority voting across all trees in the ensemble.

5.2 Algorithm Flow

1. Initialize number of trees (T), random seed, and input dataset D
2. For each tree in T:
   a. Draw a bootstrap sample from D
   b. Select a random subset R of features
   c. Construct a decision tree using Gini index on R
3. Aggregate predictions from all trees using majority vote
4. Output final classification as "Joined" or "Not Joined"



6. EXPERIMENTAL RESULTS

Prediction Results using Random Forest Algorithm

A web-based system was developed using Microsoft technologies (ASP.NET, C#, SQL Server) where HR inputs candidate data into a prediction form. The Random Forest algorithm processes the data in real-time and delivers a joining prediction.

Discussion

- The training dataset for HR recruitment was processed using the implemented Random Forest algorithm.
- The RF logic was written in custom C# code without relying on external libraries, offering full control over the behavior and integration.
- The system supports dynamic datasets, allowing predictions on real-time candidate entries.
- The algorithm achieved approximately 90% prediction accuracy, and average prediction time was recorded at ~1006 milliseconds.

.

- Results demonstrate high reliability and fast response, making the system suitable for real-time enterprise HR decision-making.

Constraint	RF Algorithm
Accuracy	90 %
Time (milli secs)	1006
Correctly Classified (precision)	90 %
Incorrectly Classified (Recall)	10 %

7. CONCLUSION

Candidate job interview process in any organization requires more time, human resources and investments. Job interview process may involve many procedures and many rounds of interview. After the interview process, candidate selection will be done and intimated. But not all selected candidates will join to the company due to many reasons. There is no process or procedure or tool or automation to find the candidates joining to the company. This motivates us to develop an automation for candidate joining prediction using efficient machine learning algorithm. Our system helps HR department in candidate joining process.

8. STATEMENTS & DECLARATIONS:

Use of AI Statement

The authors declare that they have not used generative artificial intelligence, specifically ChatGPT in the writing of this manuscript and/or in the creation of images, graphics, tables, or their corresponding captions

Conflict of Interest and Declarations:

Authorship contribution statement: Mohitha M: Carrying the Experimental work, Data curation and writing the original manuscript and original draft Meghana B S: Supervision and review of the draft.

Acknowledgements: Nil

Compliance with Ethical Standards:

Conflict of Interest : The authors state that they don't have any conflict of interest.

Animal and Human Participants: Nil

Informed consent : Authors stated that there is no informed consent in the article.

Funding : Nil

Data availability: All the data included in this research article will be provided on request

9. REFERENCES

[1]. "An Intelligent Career Guidance System using Machine Learning", Vignesh S, Shivani Priyanka C, Shree Manju H, Mythili K.: 2021

[2]. "Predicting Students' Employability using Machine Learning Approach", : Cherry D. Casuat,Enrique D. Festijo. 2020 [3]. "Application of Data Mining in Predicting College Graduates Employment", : Shouwu H,Xiaoying Li,Jia Chen., 2021

[4]. "FUTURE JOB PREDICTION IN TRIVANDRUM USING MACHINE LEARNING TECHNIQUES", Akku George Saju, 2018

[5]. "Appropriate Job Selection Using Machine Learning Techniques", Md. Ashikur Rahman Khan,Anjan Rhudra Paul ,Fardowsi Rahman,Jony Akter ,Zakia Sultana,Masudur Rahman, 2023