

# Optimizing Recruitment: Harnessing Machine Learning for Predictive Hiring Decisions

**Mukhesh Ravi**

*The Richard DeVos Graduate School  
of Management  
Northwood University  
Midland, Michigan, USA  
[ravim@northwood.edu](mailto:ravim@northwood.edu)*

**Harika Yenuga**

*The Richard DeVos Graduate School  
of Management  
Northwood University  
Midland, Michigan, USA  
[yenugah80@northwood.com](mailto:yenugah80@northwood.com)*

**Dr Itauma Itauma**

*The Richard DeVos Graduate School  
of Management  
Northwood University  
Midland, Michigan, USA  
[itaumai@northwood.com](mailto:itaumai@northwood.com)*

**Abstract—** This study investigates the application of machine learning models to optimize recruitment processes by predicting hiring decisions based on candidate profiles. The goal is to develop a predictive model that classifies candidates as “hired” or “not hired” using demographic data, qualifications, and recruitment scores. A dataset of 1,500 candidates was used, incorporating features such as age, gender, education level, work experience, and skill scores. Several machine learning algorithms, including Random Forest, Support Vector Machine, Logistic Regression, and ensemble methods (CatBoost, XGBoost), were evaluated to determine the best-performing model and identify the most influential features. Results show that CatBoost outperformed other models, achieving an accuracy of 95%, followed by Random Forest and XGBoost. Analysis of feature importance revealed that Recruitment Strategy, Education Level, and Personality Score were the top three factors influencing hiring decisions. These findings suggest that machine learning can enhance recruitment efficiency, reduce biases, and improve hiring outcomes.

**Keywords—** Machine Learning, Predictive Hiring, Recruitment, Hyperparameter Tuning, Classification & Modelling

## Introduction

The hiring process is crucial to an organization's success, as it determines the talent aligned with its strategic goals. However, traditional recruitment methods, including manual screening and interviews, are often time-consuming, resource-intensive, and prone to biases. These biases can lead to suboptimal hiring decisions, affecting diversity and organizational growth. As a result, there has been growing interest in leveraging machine learning (ML) models to improve recruitment efficiency, reduce biases, and enhance decision-making [1]. Despite their potential, existing predictive models in hiring face significant challenges, particularly in mitigating fairness issues.

## Problem Statement

Although ML models have shown promise in improving recruitment accuracy, they are susceptible to issues like dataset

imbalance, which can exacerbate biases in hiring predictions. For example, an imbalanced dataset where one gender or educational group predominates can lead to biased predictions favoring the overrepresented group, perpetuating existing disparities [2]. While existing studies have examined ensemble methods like CatBoost and XGBoost, this paper seeks to demonstrate their superiority over simpler models by highlighting their enhanced predictive performance and practical value in optimizing hiring decisions [3] [4] [5]. These models have demonstrated robust performance in handling imbalanced datasets and capturing complex, non-linear relationships, yet their application in recruitment prediction is underexplored [6].

## Objective

This research aims to develop a predictive model to improve the accuracy and fairness of hiring decisions by addressing bias-related challenges in the recruitment process. It evaluates the performance of advanced ensemble models like CatBoost and XGBoost against traditional models such as Random Forest and Logistic Regression. These advanced models are chosen for their ability to handle imbalanced datasets and capture intricate patterns in candidate data [7].

## Significance

This study contributes to the literature by providing a comparative analysis of advanced ensemble models for recruitment prediction. It emphasizes the importance of addressing dataset imbalance and bias mitigation to improve fairness in hiring decisions. By identifying key predictive features such as Recruitment Strategy, Education Level, and Personality Score, this research offers actionable insights for human resource (HR) professionals. Furthermore, it shows the potential of ML models to streamline recruitment, enhance decision accuracy, and reduce bias, thereby improving hiring outcomes and promoting diversity in the workplace.

## Research Question

The central research question guiding this study is: **Which features are most predictive of hiring decisions, and how can machine learning models be leveraged to optimize recruitment strategies while mitigating bias?** To answer this, we examine features such as education, skill scores, and work experience, with the aim of developing models that support automated, efficient, and fair hiring processes.