

Employee attrition prediction using machine learning models.

Iparraguirre-Villanueva, Orlando¹, Chauca-Huete, Luis², Prieto-Chavez, Rosas³, Paulino-Moreno, Cleoge⁴
^{1,2,3} Universidad Tecnológica del Perú, Lima, Perú, c27399@utp.edu.pe, c20903@utp.edu.pe, c20378@utp.edu.pe
⁴ Universidad Católica de Trujillo - Perú, paulinozenaida16@gmail.com

Abstract Today's business landscape is characterized by competition and dynamism, which has transformed human resource management into an essential strategic partner for organizations. Employee turnover poses risks that affect productivity and knowledge management. This study focuses on predicting employee turnover using Machine Learning (ML) models. For the training process, a dataset composed of 4410 records and 29 variables was used, in the process of training and evaluation of the ten models, the artificial intelligence (AI) method was followed. The findings showed that the XG Boost Classifier (XGBC) and Random Forest (RF) models achieved the best accuracy and performance rates, with 98.8% and 98.7%. Followed by Decision Tree Classifier (DT) with 97.6%, and the other models, such as Gradient Boosting Classifier (GBC), Ada boost Classifier (AC), Logistic Regression (LR), KN Classifier (K-NNC), SGD Classifier (SGDC), Support Vector Classifier (SVC) and Nu Support Vector Classifier (NuSVC), achieved the following rates: 88.4%, 85.4%, 84%, 82.2%, 83.0%, 83.0%, 55.0%, respectively. Finally, it is concluded that the models are useful and effective in prediction. Their practical implementation in human resource management strategies is recommended for proactive intervention.

Keywords-- Machine Learning; Artificial Intelligence; Management; Human Resources; Models.

I. INTRODUCTION

Today's business environment is competitive and dynamic, which has led to the evolution of Human Resource Management, making it a crucial strategic partner [1],[2]. Given the complexity and dynamism of external organizational environments, employee turnover is a constant risk that impacts productivity and knowledge management. Several studies reveal worrying rates of job dissatisfaction and turnover in different countries and sectors, underlining the need to address this challenge. According to the Burnout 2023 study, conducted among workers and HR specialists in five countries in the region, Argentina has the highest burnout rate (94%), followed by Chile (91%), Panama (83%), Ecuador (79%) and Peru (78%) [3]. Similarly, LinkedIn's 2021 report shows a global average turnover rate of 10.9% [4]. According to a report by the Society for Human Resource Management (SHRM) in the U.S., employee turnover can cost up to 50-60% of the replaced employee's annual salary. On average, it can take companies 42-60 days to fill a vacancy, and about 20% of new employees leave within the first 45 days [5]. In addition, a study by the Gallup consulting firm estimates that only 21% of employees

worldwide are engaged in their jobs, suggesting a general lack of job satisfaction. The International Monetary Fund (IMF) and the World Bank (WB) also estimate that global unemployment will be 5.3% and 5.2% in 2023 [6],[7]. Therefore, forward-thinking companies are turning to advanced technologies such as AI and data analytics to anticipate potential employee turnover and take preventative measures to minimize its effects [8],[9]. The advancement of technology enables proactive and timely measures to be taken in human resource management, allowing companies to deal with the threat of losing a valuable employee. This is made possible by the intrinsic potential of emerging technologies to address various issues, and they emerge as catalytic agents by enabling the generation of content from pre-existing data [10], [11]. Thus, AI is being adopted by various companies around the world for its ability to perform complex tasks that previously required human intervention. Within the field of AI, we find ML, a branch of AI that focuses on the development of algorithms and models. These models capture patterns and relationships in the data during the training process [12], [13] and are then used to make predictions or decisions about new data [14],[15], [16]. The objective of this research is to predict employee turnover using the following ML models: LR, SVC, NuSVC, XGBC, K-NNC, SGDC, DTC, RF, ABC, and GBC. In addition, the organization of this paper follows the following structure: section 2 presents related work on employee turnover. In section 3, the theoretical foundations underpinning this paper are built. Section 4 presents the training results of ML models. Section 5 discusses the results with findings from related work. Finally, Section 6 presents the conclusions and future work.

II. RELATED WORK

Several studies have successfully explored the use of ML to address employee turnover prediction. The research has highlighted the success of different models, such as Logistic Regression, ensemble models, analysis of different ML methodologies, and the use of deep models to predict employee turnover. In addition, this study builds on research that has successfully used ML to solve various problems. For example, in research [17], a model based on the Logistic Regression algorithm was presented to analyze the probability of employee attrition in an organization. The methodological approach employed consists of the exhaustive application of the Logistic Regression algorithm. This model has shown remarkable results, with an accuracy rate of 84.12%, an accuracy of 84% and a recall rate of 100%. The researchers concluded that the model has significant potential for analyzing the probability of

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

