

Performance Prediction of Employees with Data Mining Approaches

A.V. Kalpana¹, R. Dhanalakshmi^{2*}, B. Praveen Kumar³, J. Umamageswaran⁴

¹Department of Artificial Intelligence, Shri Vishnu Engineering College for Women,
Bhimavaram, Andra Pradesh, India.

²Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of
Medical and Technical Sciences, Chennai, Tamil Nadu, India.

³Department of Computer Science and Engineering, Sri Venkateswara College of Engineering,
Sriperumbudur, Tamil Nadu, India.

⁴Department of Information Technology, R.M.K Engineering College,
Chennai, Tamil Nadu, India.

*Corresponding author: dhanalakshmi.sse@saveetha.com

Abstract. To improve the skills of HR applications, numerous strategies may be applied to expand AI technology. To identify the most important patterns and rules in massive amounts of data, data mining is one of the AI technologies that has been created. Managing an organization's talent is a problem for HR professionals, especially in terms of making sure the correct individual is assigned to the right role at the right time. An outline of a few of the resource management issues that may be tackled utilising data mining approaches is presented in this article. This study also aims to apply a talent management assignment, which is to discover potential talent by using prior experience information to anticipate this performance. Employee performance data may be mined using the approaches outlined in this article. Finally, this research suggests that Data Mining Techniques might be used to predict talent in the future. Human capital is a major issue for business executives, who are primarily interested in acquiring highly qualified employees who are anticipated to perform well. The pieces are also framed using the Open-source tool's knowledge flow concept. Several trials were carried out based on the preceding methodologies to obtain a highly accurate model that can be used by decision makers and human resources experts to anticipate and improve their workers' performance.

Keywords: Prediction, Talent Assessment, Data Mining, Forecasting, Management.

INTRODUCTION

Making unstructured judgments may be made easier with the use of HR apps that incorporate Artificial Intelligence (AI). To improve the capacity of HR applications, numerous strategies may be applied to expand AI technology. To find meaningful rules and patterns in massive amounts of data, data mining is one of the AI technologies that have been created. HR data may be a valuable resource for knowledge discovery and decision-making tools. Until now, the methodologies and applications of data mining in the HRM industry have received little attention. Aiming to discover current talent by projecting the performance based on prior experience knowledge is the goal of this study [1].

Essentially, HRM encompasses a wide range of managerial activities and duties related to workforce development and management. Individual and organisational competitiveness, as well as the enhancement of productivity, quality, and compliance with legal or social obligations, are the primary goals of human resources management (HRM) [2]. In addition, each firm must effectively compete in terms of pricing, quality, service, or innovation. When you have the proper personnel in place, you'll be able to do all of these things and more. Recent concerns for HR managers include ensuring the appropriate person is in the right place at the right time when it comes to talent management [3].

These duties need a great deal of managerial decision-making, which may be challenging and ambiguous, at times [4]. Human experience, knowledge, preference, and judgement all play a role in modern HR decision-

Received: 09.02.2022 Revised: 31.03.2022 Accepted: 14.04.2022
Licensed under a CC-BY 4.0 license | Copyright (c) by the authors

making methods. Inconsistent, imprecise, uneven, and unexpected judgments might be the result of several circumstances. As a result, this condition might lead to feelings of injustice when it comes to supporting individual growth and development [5]. In addition, the production of the organisation may be affected in the future as a result of this. Overview of Data mining process is shown in Figure 1.

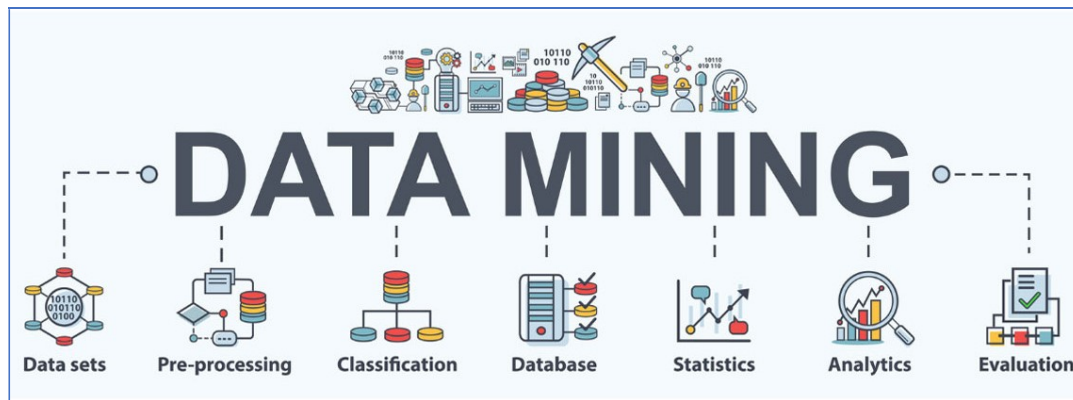


FIGURE 1. Overview of Data Mining Process

One of the most difficult aspects of talent management is determining which employees have the potential to succeed, and this may be accomplished utilising the Data Mining approach [6]. As a result, the goal of this research is to provide insight into prospective Data Mining methodologies for predicting talent. HR decision support systems can include these possible strategies to create a system that aids managers in selecting the right individual for the right role [7].

EXISTING WORK

There is a lot of attention being paid to data mining (DM) as a novel method of data analysis. As of late, data mining has received a lot of attention and worry from both the information business and the general public alike. For the simple reason that so much data can be easily accessed, as well as the necessity to transform that data into usable information and knowledge [8]. A Decision Support System (DSS) interfaced with a DM tool can assist executives in making better informed and objective judgments.

In addition, it may assist managers in retrieving, summarising, and analysing data linked to decision-making so that better judgments can be made. Clustering, association, classification, and prediction are all types of data mining challenges. There have been several developments in Data Mining over the years including statistical analysis as well as use of neural networks and decision trees as well as genetic algorithms. As well as these applications, Data mining has been used in a wide range of other industries as well [9].

Despite this, it is an uncommon occurrence in HRM. It is now possible to solve HRM issues using the Data mining technique. This method is frequently employed in the recruitment process in order to narrow down the pool of potential employees. However, there have been a few examples of predictive applications in human resources management (HRM), including sales premiums and the persistence indices for insurance agents, as well as the mis-operation behaviours of operators [10]. Therefore, in this study, data mining techniques are used to predict future workers as part of the talent management process is applied. Talent management has become an increasingly important approach to HR activities in every firm. Anyone with the potential to have a substantial impact on the organisation's current as well as future performance is regarded to have talent [11].

There are several aspects to human resource planning that have to do with how a firm manages its employees [12]. As well as making certain that the right people are in the right jobs, performance appraisal can be defined as a procedure for maintaining involved in maintaining in important positions and encouraging individual advancement, as well as a decision to manage the supply, demand, and flow of natural ability through an organization's "human capital engine." [13].

PROPOSED SYSTEM

Most of the Data Mining research in the HR area focuses on staff selection, whereas just a few studies relate to other tasks including planning, training, and talent management [14]. HRM has recently sought a strategic position by using Data Mining methodologies considering the changing expectations and increasing visibility of HR management [15]. This may be accomplished by the identification of relevant knowledge created from existing patterns in HR databases. A person's aptitude in this research is looked into for trends that indicate. Data mining techniques may be used to create the patterns.

To find the most appropriate Data Mining methods, it is critical to connect data mining difficulties with talent management requirements. Use of C4.5 and Random Forest for decision trees; MLP and Radial Basic Function Network for neural networks will be tested in this project. First, the classifier algorithms selected for this study on a sample of employee data is run. In this example, a classifier for HR data that's accurate in terms of its methods is investigated. The proposed system architecture is shown in Figure 2 [16].

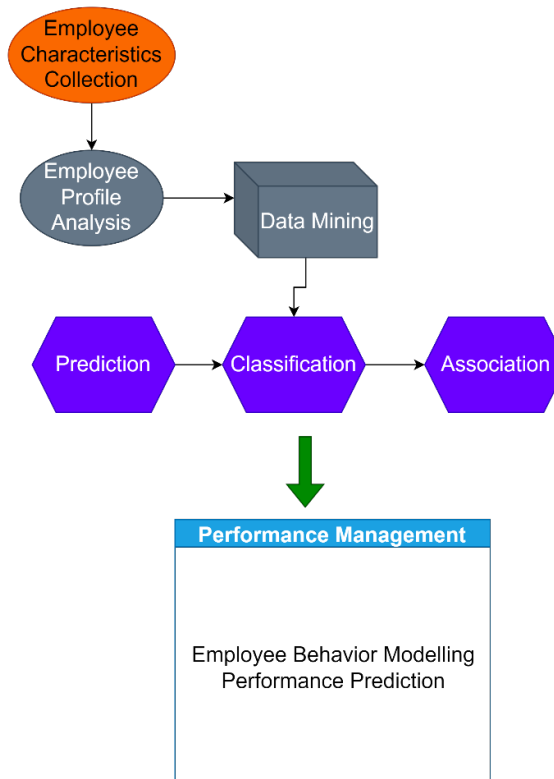


FIGURE 2. Proposed System Architecture

The accuracy of these two data mining approaches was found to be better than 70% in the early studies, which can serve as an indicator for the best Data Mining methods for HR data. To test the accuracy of common data mining prediction algorithms, this experiment is conducted as said earlier. Other classifier algorithms, such as SVM (support vector machine), Rough Set Theory, Fuzzy clustering, and many more, can be used to support the accuracy of these approaches to identify the best Data Mining strategies.

One of the most often used data mining techniques is association. By looking at the relationship between one item and other goods in the same transaction, a pattern may be established. As an example, in market basket analysis, the association approach is used to find which goods people usually purchase together. Businesses may use this information to create targeted marketing campaigns that will increase sales and, ultimately, profits. Data mining techniques like association are well-known.

By looking at the relationship between one item and other goods in the same transaction, a pattern may be established. As an example, in market basket analysis, the association approach is used to determine which goods people usually purchase together. Businesses may use this information to create targeted marketing campaigns that will increase sales and, ultimately, profits. The findings suggest that decision trees and neural networks are both effective tools for predicting future talent needs and performance.

RESULTS AND DISCUSSION

In this experimental study, turnover may be predicted by factors like as absenteeism and tardiness as well as work content, demography, and previous team experience. These findings can be interpreted in several ways, depending on the exact situation. Job content factors such as job type and subject knowledge showed up as important variables in the models. This variable has two aspects. In Table 1, various attributes considered are listed.

TABLE 1. Attribute based on Factors.

Factor	Attributes
Basic Attributes	6
Performance	18
Skill set	12
Managerial Ability	2
Personality	9

Global clients in a range of industries rely on India's software services industry to help them run the businesses. It also builds and manages proprietary and standard systems software. Software development centres in India are being used by a growing number of product firms to produce a large portion of the goods.

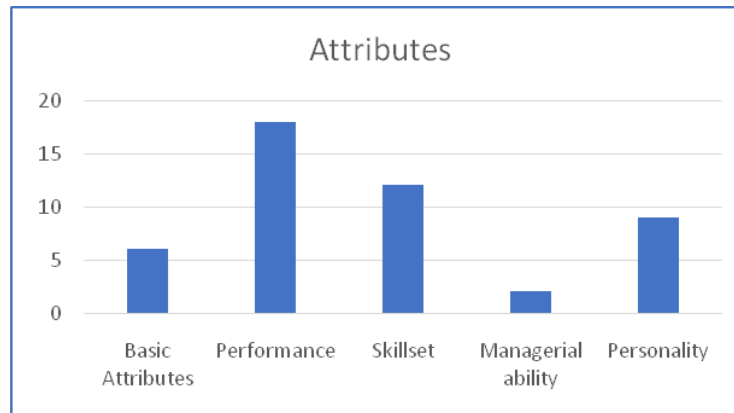


FIGURE 3. Dataset Attributes

As a result, many significant Indian and global corporations are involved in a wide range of industries and technology. Many young engineers think that gaining expertise on a wide range of platforms and domains is essential. Now, specialised knowledge is undervalued. It is feasible that a company with a superior domain or technology platform can draw employees from those in organisations with less favourable platforms, because of the need to operate across several platforms or domains. Table 2 lists the outcome of accuracy analysis.

TABLE 2. Accuracy Analysis

Train /Test	Accuracy
90:10	64.25
80:20	73.86
70:30	75.96
60:40	75.26
50:50	81.23

New projects are given to personnel who have worked on previous ones, or who are currently on the bench. There's a chance that these workers aren't too fond of the new technology. Employees may want to leave a position if it forces them to operate with an outdated technology. After a few months or years of work experience, it's possible that people start looking for new challenges and will quit the current firm if this isn't provided with such difficulties. This might explain why these findings were made. When it comes to employee turnover, age and marital status have a major role.

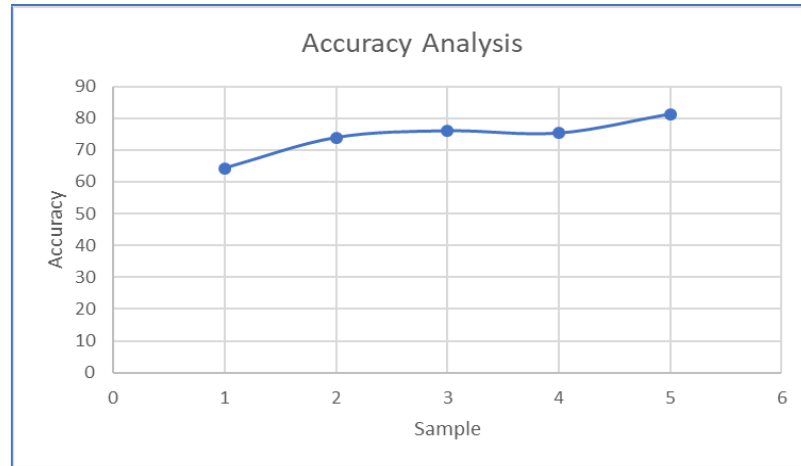


FIGURE 4. Comparing obtained Accuracy Values

The data reveal that younger employees are more prone to turnover than older employees in early stages. More than half of the young people in the outsourcing industry had changed employment in the last three years according to Hill and Associates' assessment on voluntary turnover rates. As a result, the IT industry has been able to attract a huge number of recent graduates. Figure 4 compares the obtained accuracy values from the experiment.

Graduating students who have never worked before tending to have high expectations of both them and the employers. It is especially prone to staff turnover because of the inflated expectations, which are compounded by a dearth of available job skills and skyrocketing wages. During the first few jobs, it's very probable that many individuals will go on a journey of professional discovery.

CONCLUSION

The relevance of the study, the foundation of information retrieval, data mining in human resource applications, and an overview of talent management have been discussed in this article. There are many distinct types of HR applications that scholars have explored in the literature. Nevertheless, additional HR applications and Data Mining techniques should be applied to other issue domains in HRM field research to widen the horizons of academic and practical work on HR applications utilising Data Mining methods. This is why this recommends data mining approaches based on an initial trial for performance prediction. The correctness of the methods for talent management data may be evaluated using additional Data Mining approaches for future work. It's also important to analyse the relevance of characteristics to the classifier's accuracy. This experiment should be followed up by an attribute reduction experiment to identify the most important characteristics associated with each element. The following modelling stages may be developed after the required attributes are in place. To sum it all up, the power of HR application is in the ability to adapt and gain new insights on a constant basis.

REFERENCES

- [1]. C. F. Chien and L. F. Chen, 2007, "Using rough set theory to recruit and retain high-potential talents for semiconductor manufacturing," *IEEE Trans. on Semiconductor Manufacturing*, **20(4)** pp. 528-541.

- [2]. K. K. Chen, M. Y. Chen, H. J. Wu and Y. L. Lee, 2007, "Constructing a web-based employee training expert system with data mining approach," *The 9th IEEE Int. Conf. on E-Commerce Tech. and the 4th IEEE Int. Conf. on Enterprise Computing, E-Commerce and E-Services (CEC-EEE 2007)*, pp. 659-664.
- [3]. C. F. Chien and L. F. Chen, 2008, "Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry," *Expert Systems with applications*, **34(1)** pp. 280-290.
- [4]. A. S. DeNisi and R. W. Griffin, 2005, "Human resource management," *Dreamtech Press*.
- [5]. J. Han, J. Pei and M. Kamber, 2011, "Data mining: concepts and techniques," *Elsevier*.
- [6]. M. J. Huang, Y. L. Tsou and S. C. Lee 2006, "Integrating fuzzy data mining and fuzzy artificial neural networks for discovering implicit knowledge," *Knowledge-Based Systems*, **19(6)** pp. 396-403.
- [7]. L. Morton, "Talent management value imperatives: Strategies for execution," *Conference Board*.
- [8]. N. Rathi, Human resource challenges in Indian software industry: An empirical study of employee turnover, *Int. J. Human Resources Development and Management*, **3(3)**, pp. 249-264.
- [9]. W. S. Tai and C. C. Hsu, 2006, "A realistic personnel selection tool based on fuzzy data mining method," *9th Joint Int. Conf. on Information Sciences (JCIS-06)*, pp. 190-193. Atlantis Press.
- [10]. G. K. Tso and K. K. Yau, 2007, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, **32(9)** pp. 1761-1768.
- [11]. K. Y. Tung, C. Huang, S. L. Chen and C. T. Shih, 2005 "Mining the Generation Xers' job attitudes by artificial neural network and decision tree—empirical evidence in Taiwan," *Expert Systems with Applications*, **29(4)** pp. 783-794.
- [12]. I. Cunningham, 2007, "Talent management: making it real," *Development and Learning in Organizations: An Int. J.*, **21(2)**, pp. 4-6.
- [13]. R. Mano-Negrin, 2001, "An occupational preference model of turnover behaviour-The case of Israel's medical sector employees," *J. of Management in Medicine*, pp.1-24.
- [14]. K. Morrell, J. Loan-Clarke and A. Wilkinson, 2004, "The role of shocks in employee turnover," *British J. of Management*, **15(4)** pp. 335-349.
- [15]. R. Santhosh and M. Mohanapriya, 2021, "Generalized fuzzy logic based performance prediction in data mining," *Materials Today: Proc.*, **45** pp. 1770-1774.
- [16]. S Murugan, S. Mohan Kumar and T.R. Ganesh Babu, 2020, "Convolutional Neural Network-based MRI Brain tumor classification system," *Int. J. of MC Square Sci. Res.* **12(3)**, pp. 1-8.