IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005090|

Employee's Performance Analysis and Prediction Using Random Forest Algorithm

Madhura Pradeep Banarase, Sarvesh Omprakash Sharma, Bhagyashri Bhagawan Giri, Tushar Machhindra Kalaskar, Prof. R.A. Meshram

UG Student, Dept. of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

UG Student, Dept. of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

UG Student, Dept. of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

UG Student, Dept. of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

Assistant Professor, Dept. of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

ABSTRACT: The employee data are evaluated for giving promotion, yearly increment and career advancement. Main objective is to evaluate the performance of employee. In this project we will concentrate on collecting data about employees, generating an algorithm from historical data, testing the algorithm with attributes of an employee and generating the output as whether to give the promotion or not. The information about an employee is collected by using the user interface. This information is compared with the trained data stored in the algorithm. The final goal node is to determine whether the employee will get yearly increment, promotion or not.

KEYWORDS: Employee Performance Evaluation; Employee Performance Analysis; Performance Prediction; Random Forest Algorithm; Employee Promotion.

1. INTRODUCTION

Most organizations or companies have a formal performance appraisal system where staff performance is scheduled regularly, usually once or twice a year. A performance appraisal system can greatly benefit an organization. It assists the employee in conduct within the organization's objectives by allowing employees to know what is expected of them, and provides information on employment decisions, such as those relating to salary, promotions or dismissals. An employee can improve their performance by monitoring the progress of their work. Machine learning algorithms that is, the decision tree of the data mining method can be used.

The decision tree is used to make the right decision for the Project. Based on the results of the job interview you may decide whether further training, talent enrichment or other qualifications are required or not. With the growth of staff engagement, the results of the work done have been very satisfactory, in line with Pang and Lu's theory, the organization will function better due to the high staff engagement that has been built on the individual [1]. Findings [2] mean that with the creation of high employee engagement, performance will increase, then working people will voluntarily provide additional service to their customers. Similarly, the findings [3] mean that employee interaction can improve individual performance, and people who feel busy in their work will be encouraged to provide additional services from their core business to clients.

The current system for determining staff performance is a manual. Staff performance appraisal is entirely dependent on senior management. If the authorities are biased, the employee will never be promoted. Therefore, to make this process automated, we have developed a system to monitor the performance of employees online. In order to

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005090|

provide for the annual increase of the employee, it must be assessed using the previous employee history data, which leads to an employee's fair and honest progress.

II. LITERATURE SURVEY

In recent years, some statistical analysis methods have been used to predict or analyse the purpose of employee benefits. Chien [6] used a two-stage analytical approach to predict the purpose of employee benefits. Wu [7] used the decision tree to predict the purpose of employee benefits.[8] initiated a focused profit-making approach with Holtom et al. [9] improved social networking metrics for oscillation and intermediate response time to identify changes in the communication behaviour of managers who are about to resign.

Labour departments need to identify individual employee characteristics, workplace characteristics, and work attitudes that may be related to employee benefits. Jung and Yoon [10] surveyed the staff at deluxe hotels and found that job satisfaction was related to the profit motive. Labrague et al. [11] analysed the benefits of nurses in the Philippines and found that age, job satisfaction, and job stress were factors that had a significant impact on profitability. Tran [12] investigated the motives for the transfer of high school principals and found that salary satisfaction and school achievements were negatively correlated with the goal of making a profit.

In the findings [6], employee interactions are also able to mediate relationships between variables and may reduce the occurrence of a benefit to the individual rather than the performance of the activity. From the findings of these findings, it has been proven that high employee engagement can have a positive and positive impact on job performance [4], based on strengthening previous theory. Findings [3] that people who feel busy with their work and have a high sense of self and involvement in activities in their organization will produce good work. This is in line with existing findings that reflect the interaction of employees, directly and indirectly, affecting individual performance [5], can lead to fatigue in the pursuit of profit [4]. It has been found [6] those positive emotions can promote employee interaction to improve work efficiency through better service delivery, yes [4] of decision. The K integration algorithm is a common way to define functional classes. The author uses the K algorithm integration algorithm to Employee into a separate collection based on their Working Quality. The author has used the decision tree algorithm to distinguish the employee easily and make the right decision quickly.

III. METHODOLOGY

3.1) Random Forests

WQRF implementation is based on a random forest algorithm (RF). RF is a combination of the algorithm proposed by Breiman [13] in 2001 where if the predicted result is a different value, it is a random forest system, and if it is a continuous value, it is a random forest retreat. The RF separation algorithm is used in two phases. First, the RF algorithm extracts small samples from the original samples using a bootstrap sample duplicate method and creates trees for decision samples for each sample. Second, the algorithm divides the decision trees and uses a simple vote, with the largest split vote as the final predictor result.

Based on employee benefit data with non-standard features with high magnitude, this study prioritizes the WQRF algorithm, which uses the RF algorithm in two phases. First, an algorithm is used to measure the value of each element and to reduce the size. Second, selected features are used with the RF algorithm and performance test results are calculated for the F value of each tree to determine the different tree weights to produce a predictive model.

3.2) Classifier Evaluation Index

Common indicators for evaluating the performance of a precision prediction model (ACC), memory, accuracy (PPV), and curve point (AUC). To calculate these indices, a confusion matrix is used. In the matrix, the columns represent the prediction categories and the total value in the column is the view of the data in the category. In addition, the lines in the matrix represent the actual categories and the total value in the lines represents the visibility of the data in that category. In this study, we focused on whether there is an employee benefit or not, which is a binary system.

AUC refers to the area below the curve of the receiver operating curve (ROC). It is an important indicator of the advantages and disadvantages of a binary prediction model. The greater its value, the better the performance of the model.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005090|

3.3) Features Ranking

There are many aspects of employees within an organization and can be divided into several common categories: basic knowledge, job knowledge, job position knowledge, academic knowledge, training knowledge, degree certificate, etc.If all the factors (size) are considered, this will lead to significant costs in terms of time and space in the implementation of the algorithm, which has a significant impact on its performance. Therefore, the RF method is first used to reduce the size.The RF method can measure features (variables) based on their value, so it can be used to reduce the size and subtract less important features. The main idea is to calculate the decrease in the accuracy of the RF prediction by adding sound to each element.

The objectives of feature selection are:

1) Improve model prediction accuracy.
2) Creating a fast and low utility model, as well
3) Make the model more descriptive.

This study uses the RF method to measure employee characteristics and to reduce the size of employee benefit forecasting. Maintaining the m of the most important aspects of employee benefit will increase the efficiency of forecasting.

3.4) Weighted Random Forest Algorithm

Generally, all RF decision trees have the same weight value when voting for division. However, it has a dangerous feature when used with an unpredictable data segmentation forecast. To address this, we introduced a weighted F ratio to the RF algorithm, which produces better profit-predictive performance by assigning different weights to trees of different resolutions. From a data mining perspective, the problem in employee benefit prediction is the fragmentation of unbalanced data.

We classify "profit" as a positive category, while "leave" as a negative category. Obviously, the good category is small and the bad category is the largest. It is not enough to measure the performance of a model accurately with unequal data; for example, if the company's staff turnover rate is 2% and no one is expected to leave, the accuracy level may be as high as 98%; however, this does not make sense here.

In this study, it is better not to misunderstand an employee who is unwilling to leave as he or she may be leaving than to ignore a person with a sincere intention to leave. This study incorporates two diagnostic criteria for accuracy and memory and uses the harmonic definition of F-measure to evaluate the effectiveness of each decision tree and to calculate the weight of the vote. Compared to the standard RF algorithm, this updated algorithm improves the performance of unbalanced data splitting.

3.5) WQRF Method Overview

The WQRF algorithm is proposed to build a predictive model. Original (historical) data is broken down into a training set and verification is set randomly. In addition, unselected data (OOB) is used as a test set. In the first calculation, a training set is used to measure features on the importance of using the RF algorithm. Weights in a weighted RF algorithm are obtained using a confirmation set to calculate the F value of each tree. In any employee, the features of m are obtained and incorporated into the guessing model, so that the employee's intention to move in the future is not predicted. The prediction model can be tested in the form of 10-fold cross-sectional verification for accuracy, sensitivity, accuracy, and AUC.

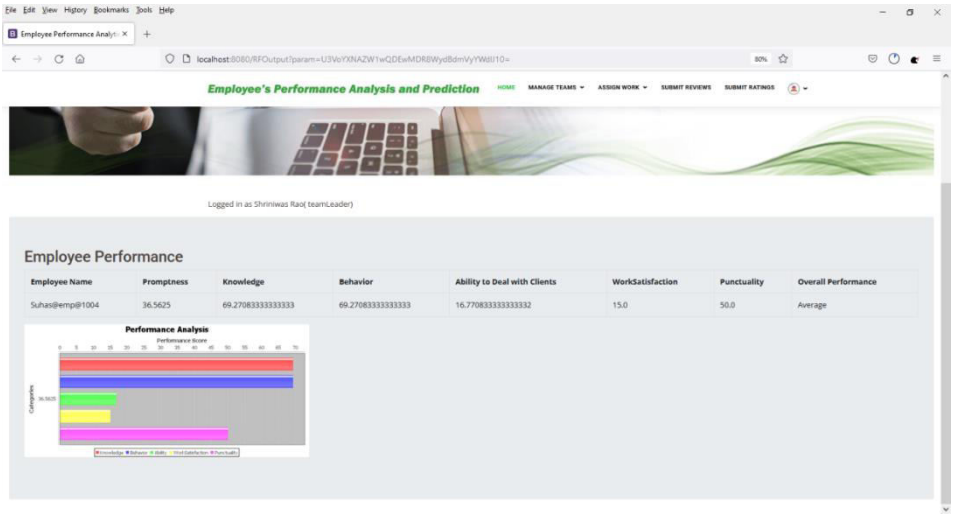
IV. RESULTS

Using the algorithms and technologies as stated in this study, an organization can evaluate the performance of its employees based on the pre-defined deciding factors.A representation of the performance evaluation can be seen as shown in the following screenshot.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005090|



Hence, it becomes easier for higher authorities to classify their employees according to their abilities and deciding promotions becomes hassle free.

V. CONCLUSION AND FUTURE WORK

In this study, an improved RF algorithm, WQRF based on F-weighted scale, was proposed. The main idea is to follow two steps. First, a random forest algorithm is used to order feature value and reduce size. Second, the selected features are used with the random forest algorithm and the F-value values are calculated on each decision tree as the weights for building a predictive model. In the employee database of the telecommunications company'sbranch in China, key factors influencing employee profits are identified using the WQRF algorithm.The algorithm proposed here is also used to predict employee income in industries such as education, medicine, finance, and other fields. If an organization can predict in advance which employees will be most profitable, it can create a system and take steps to minimize the risk, address the need for immediate replacements, and make adjustments to keep staff in key positions. Using the WQRF method, HR managers can predict better employee benefits and take action in a timely manner. The Tis algorithm can also be applied to the problem of unequal data allocation of other fields, for example, to customer rhetoric, cancer screening, and randomized testing. These problems have similarities, such as focusing on a small group of people, and the requirement that “it is better to test than to see.” Compared to other algorithms, in general, the WQRF algorithm has a higher memory value for a smaller population. In the future, we will theoretically analyse data features in these fields and further validate and prepare for the WQRF. In terms of research limitations, adding a filter element and weighting procedures means that modelling has a higher cost over time than other algorithms. In future studies, improvements in efficiency and overall accuracy of forecasting may be analysed. In addition, this algorithm is designed for unequal data and is not suitable for high-profit industries. How to expand the universal algorithm still needs to be studied further.

REFERENCES

[1] A. Denvir and F. McMahon, “Labor turnover in London hotels and the cost effectiveness of preventative measures,” International Journal of Hospitality Management, vol. 11, no. 2, pp. 143–154, 1992.

[2] T. H. Feeley, J. Hwang, and G. A. Barnett, “Predicting employee turnover from friendship networks,” Journal of Applied Communication Research, vol. 36, no. 1, pp. 56–73, 2008.

[3] P. A. Gloor, A. Fronzetti Colladon, F. Grippa, and G. Giacomelli, “Forecasting managerial turnover through e-mail based social network analysis,” Computers in Human Behavior, vol. 71, pp. 343–352, 2017.

[4] W. C. Hong, P.-F. Pai, -Y. Huang, and -L. Yang, “Application of support vector machines in predicting employee turnover based on job performance,” in Proceedings of the International Conference on Advances in Natural Computation, pp. 668–674, Springer, Berlin, Germany, 2005.

[5] H.-W. Kao, S.-W. Lin, and S.-Y. Wan, “Applying decision tree to predict nursing turnover-a case study in a public hospital,” The Journal of Taiwan Association for Medical Informatics, vol. 21, no. 4, pp. 15– 29, 2012.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005090|

[6] H.-J. Chien, Application of the two-stage cluster analysis on employee voluntary turnover intention [M.S. thesis], Yuan Ze University, 2007.

[7] C.-T. Wu, Using the decision tree approach to forecast contractor turn over trend – a case study of wafer foundry [M.S. thesis], National Cheng Kung University, 2008.

[8] T. R. Mitchell, B. C. Holtom, T. W. Lee, C. J. Sablynski, and M. Erez, "Why people stay: using job embeddedness to predict voluntary turnover," Academy of Management Journal, vol. 44, no. 6, pp. 1102–1121, 2001.

[9] B. C. Holtom, D. R. Smith, D. R. Lindsay, and J. P. Burton, "Te relative strength of job attitudes and job embeddedness in predicting turnover in a U.S. military academy," Military Psychology, vol. 26, no. 5-6, pp. 397–408, 2017.

[10] H. S. Jung and H. H. Yoon, "Error management culture and turnover intent among food and beverage employees in deluxe hotels: the mediating effect of job satisfaction," Service Business, vol. 11, no. 4, pp. 785–802, 2017.

[11] L. J. Labrague, D. Gloe, D. M. McEnroe, K. Konstantinos, and P. Colet, "Factors influencing turnover intention among registered nurses in Samar Philippines," Applied Nursing Research, vol. 39, pp. 200–206, 2018.

[12] H. Tran, "The impact of pay satisfaction and school achievement on high school principals' turnover intentions," Educational Management Administration and Leadership, vol. 45, no. 4, pp. 279–290, 2016.

[13] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[14] S. Malek, R. Gunalan, S. Kedija et al., "Random Forest and Self Organizing Maps application for analysis of pediatric fracture healing time of the lower limb," Neurocomputing, vol. 272, pp. 55–62, 2018.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details