

Employee Performance Prediction Analysis using Random Tree Classification Technique

Santhipriya Gunturu¹ Mr. G. Ananthnath²

¹Student ²Assistant Professor

^{1,2}Department of Computer Applications

^{1,2}KMM Institute of PG Studies, Tirupati, India

Abstract— Almost every corporation has a performance management system or conduct a performance analysis of its employees. Though there is little evidence that this existing system is able for evaluating the performance of the worker in a big company. The intend of this paper is set references and guidelines for performance analysis of employees for a typical sales company. Our system takes in employee data such as the attendance, pay rate, revenue, expense, gender, and remark to evaluate the performance of the employees of the corporation. Our system would be using Random tree Classification algorithm for evaluate the performance of the employees.

Key words: Data-Mining, Employee Performance Analysis, Data Mining Classification

I. INTRODUCTION

A tremendous growth in the available data as a result of the computerization of the every aspect of organization has immensely to the development of the intelligent decision making technologies. A young but promising of these kind of technology is Data Mining which is the process of analyzing data from different perspectives and summarizing it into useful information. Data Mining aims at developing knowledge from available data and could further be used for improving the process as the whole. There are many classification techniques in Data Mining such as Decision Tree, Neural network, irregular Set Theory, Bayesian theory and Fuzzy logic. In this we have considered the Bayesian theory for the purpose of classification. Bayesian Classification is less complex as compared to the other classification but does affect the accuracy of the system due its less complexity. Data mining stand out due to its wide array of techniques from the different domains such as statistics, artificial intelligence, machine learning, database systems and visualization. These serve as groundwork for its application to employee performance analysis for a sales company. We basically use the training data with the five components in it to train the system to construct the Bayesian Table. Based on which a score is calculated which further being used to generate the performance rating of each employee. Using Bayesian Classification algorithm we are able to calculate the score of the employee. The leave to a specific employee granted depending on the performance score of the employee. In the end we finally try to illustrate the performance of the specific employee graphically in comparison to other employees. This graphical depiction typically illustrates the performance of an employee in comparison to other employees with the help of the bar chart. Decision tree is one of the most popular machine learning algorithms used all along, Decision trees are used for classification and regression problems, this story we talk about classification. Decision tree is the main data mining

tool used to build the classification model, where several classification rules were generated. The Bayesian Classification represents a supervised learning method as well as a statistical method for classification. Assumes a primary probabilistic model and it allows us to capture ambiguity about the model in a just way by determining probabilities of the outcomes.

II. RELATED WORK

A. Application of Data Mining Classification in Employee Performance Prediction

In emerging knowledge economy such as Kenya, organizations rely heavily on their human capital to build value. Consequently, performance management at the individual employee level is critical and the business case for implementing a system to measure and improve employee performance is strong. Data Mining can be used for knowledge discovery of attention in Human Resources Management (HRM). We used the Data Mining classification technique for the extraction of knowledge significant for predicting employee performance using previous assessment records a public management development institute in Kenya. The Cross Industry normal Process for Data Mining (CRISP-DM) was adopted for predictive analysis. Decision tree was the main Data Mining tool used to make the classification model, where several classification rules were generated. To validate the developed model, a prototype was constructed and the data collected from the institute's Human Resource branch was used. Results show that employee performance was extremely affected by experience, age, academic qualification, professional training, gender, marital status and previous performance appraisal scores. This paper propose a prediction model for employee performance forecasting that enables the human resource professionals to refocus on human capability criteria and thereby enhance the performance appraisal process of its human capital.

B. Data mining: Evaluating Performance of Employee's using Classification Algorithm Based on Random Tree.

The main objective is to evaluate the performance of employee using Decision Tree algorithm. The data mining classification method like decision tree, rule mining, clustering etc. can be useful for predicting the performance of an employee working in an organization. The employee data are evaluate for giving support, yearly increment and career advancement. In order to provide yearly increment for an employee, it should be evaluate by using past historical data of employees. The historical data store in the table are subjected to learning by using the decision tree algorithm and the performance are found by testing the attribute of an employee against the rules generate by the decision tree classifier. This paper concentrates on collect data about

employees, generating a decision tree from the historical data, testing the decision tree with attributes of an employee and generating the output as whether to give the promotion or not. The information about an employee are collected by using the user line. This information is compared with the trained data store in the decision tree. The final goal node is to settle on whether the employee will get yearly increment, promotion or not.

C. Domain driven Data Mining in Human Resource Management

An increasing number of publications relating to data mining in the theme of human source management (HRM) indicate the presence of a prospering new research field. The current paper reviews this research on HR data mining to analytically uncover recent advancements and suggest areas for future work. Based on the come up to of domain driven data mining, an initial framework with significant domain-specific necessities is elaborated. Relevant research donations are identified and reviewed against the background of this framework. The reviews reveal that HRM constitute a noteworthy new domain of data mining research that is subject by method- and technology-oriented work. However, specific domain requirements, such as evaluate the domain success or complying with legal standards, are frequently not recognized or considered in existing research. Therefore, the organized consideration of domain-specific food is demonstrated here to have significant implications for outlook research on data mining in HRM.

III. PROPOSED ALGORITHM

A. Random Tree Algorithm:

Random tree is one of the mainly popular machine learning algorithms used all beside, random trees are old for classification and regression problems, this account we talk about classification. Random tree is the main data mining tool used to make the classification model, where several classification rules were generate. Used a Naive Bayes classifier to predict job performance in a call hub with the aim of knowing what levels of the attributes are indicative of individuals who execute well. By using set records, they predicted future performance of sales agents, achieving suitable results. Random tree and association rules to make useful rules for personnel selection. This structure can be used to develop an useful personnel selection mechanism to find the talents who are the most suitable to their own organizations.

B. Navie Bayesian Algorithm Classification

The Bayesian Classification represent a supervised learning method as well as a statistical method for classification. assume an underlying probabilistic model and it allow us to capture uncertainty about the model in a principled way by determining probabilities of the outcomes. It can answer diagnostic and predictive problems. Bayesian classification provides realistic learning algorithms and prior knowledge and practical data can be combined. Bayesian Classification provides a helpful perspective for kind and evaluating many learning algorithms. It calculate explicit probabilities for premise and it is robust to noise in input data.

IV. RESULT

The employee details are uploading each employee data enter person experiences, qualification based on performance are analysis each particular data set are graphical represent in below screens more employee performances experiences and each person data every day search performances increment.

A. View

EmpId	Qualification	Age	Gender	Marital Status	Experience	Training Period
934P7Z798T	SSC	24	Male	No	1	Yes
517Y3C346F	MBA	31	Female	No	3	Yes
996A9C924B	Degree	30	Male	No	2	Yes
941Z5W533C	BCom	20	Female	Yes	6	Yes
779Q2V56X	Inter	23	Female	Yes	7	Yes
78KCF723L	Degree	27	Male	No	0	Yes
234A6W53F	MBA	22	Male	No	0	No
368B7C750G	Degree	22	Male	Yes	2	Yes

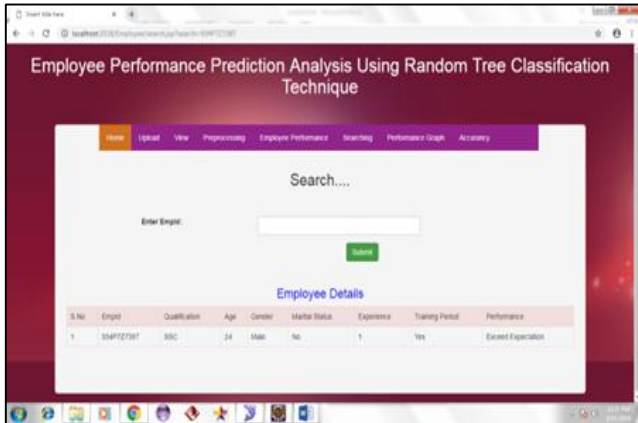
B. Preprocessing

EmpId	Performance
934P7Z798T	Exceed Expectation
517Y3C346F	Does Not Meet Minimum
996A9C924B	Need Improvement
941Z5W533C	Does Not Meet Minimum
779Q2V56X	Does Not Meet Minimum
78KCF723L	Meet Expectation
234A6W53F	Exceed Expectation
368B7C750G	Need Improvement

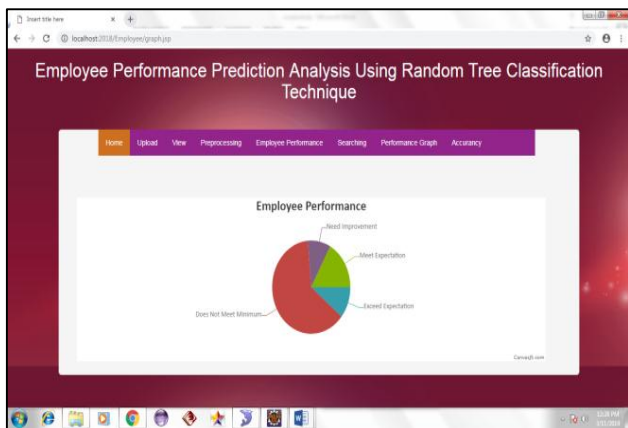
C. Employee Performance

EmpId	Performance
934P7Z798T	Exceed Expectation
517Y3C346F	Does Not Meet Minimum
996A9C924B	Need Improvement
941Z5W533C	Does Not Meet Minimum
779Q2V56X	Does Not Meet Minimum
78KCF723L	Meet Expectation
234A6W53F	Exceed Expectation
368B7C750G	Need Improvement

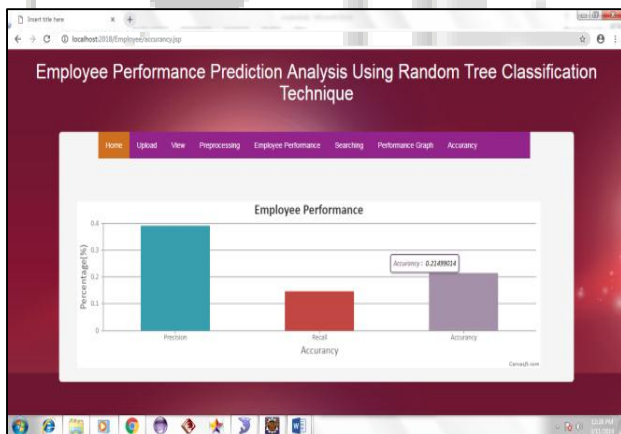
D. Search



E. Graph



F. Accuracy



V. CONCLUSION

This paper focused on the option of building a classification model for predicting employee performance. We used 5 attributes for study of the performance of the employees. These 5 attribute were used to estimate the performance score. The classifier used in the learn was the Random tree Classifier. Our system is helpful for big company which judges its employees on normal basis based on the revenue, pay rate, attendance, employee as an expense to company and performance of the employee. A better system than our accessible model could be implemented with the algorithm like Random tree algorithm which could work on additional complex employee data and with better accuracy.

REFERENCES

- [1] Witten, I., H., Frank, E., (2005). Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers, Elsevier Inc.
- [2] Mishra, P., Padhy, N., Panigrahi, R. (2013). The survey of Data Mining applications and feature scope, Asian Journal of Computer Science & Information Technology, 2(4), pp 67 – 77.
- [3] John M. Kirimi, Christopher A. Moturi (2016). Application of Data Mining Classification in Employee Performance Prediction, International Journal of Computer Applications (0975 – 8887) Volume 146 – No.7, July 2016
- [4] Ashish Dutt, Maizatul, Akmar Ismail (2016), A systematic review on Educational Data Mining. IEEE (2016).
- [5] RohitKotalwar, Raj Chavan, Sunny Gandhi, VinayakParmar (2014),Data mining: Evaluating Performance of Employee's using Classification Algorithm Based on Decision Tree. IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498 Vol.4, No. 2, April 2014.
- [6] Tom Shott, Chris Imondi, Ryan Fedie (2012), Evaluation of Performance Management Systems for Knowledge Workers,2012 Proceedings of PICMET '12: Technology Management for Emerging Technologies.
- [7] Qasem A. Al-Radaideh, Eman Al Nagi (2012), Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance,(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No. 2, 2012.