Employee Performance Evaluation System Using ID3 Algorithm

Shubham Tupe[1], Chetan Mahajan[2], Dnyaneshwar Uplenchwar[3], Pratik Deo[4]

B.E, Department of Computer Engineering, SKNCOE. Savitribai Phule Pune University, Pune,

Maharashtra, India[1,2,3,4]

ABSTRACT: In industrial sector employee performance evaluation is very important field which will be used for predict the performance of employee. By considering kinds of fields like Employee attendance, projects, skills, work quality, Behavior ,Creativity etc. Data mining consist of different techniques which can be used to complete this goal. From all this data we need to find the data which will give more information which will help to predict the performance. Data Mining techniques allows us to discover an hidden patterns, relationship form huge amount of data. Information extracted from large database is helpful in decision making. By referring the extracted information the processing methodologies are selected. The performance of employee is calculated in order to know them well and decide its position in organization.

KEYWORDS: Data Mining, Decision Making, Different Standards, Employee Performance Evaluation.

I.INTRODUCTION

Data mining is often called as knowledge discovery Procedure. We used data mining techniques in order to identify the patterns and relationships among them.Data mining consist of combinations of Machine learning, Statics, Visualization for identifying the patterns.Here we uses an decision tree for classifying data into variable types of classes which are useful for prediction purpose. This kind of technique is very beneficial in order to predict the performance of an employee very accurately by considering the decision tree.To improve the performance of an employee supervisor monitors the data very carefully.

II. LITERATURE SURVEY OR EXISTING SYSTEM

In previous system employee performance is calculated by using some parameter like Attendance, Quality of work, creativity. By considering parameters an entropy is calculated in order to calculate an performance of an employee. In previous system information gain and entropy plays important role in order to make an prediction about employee performance. In existing system decision tree was used in order to make a decision by considering an information gain an entropy. In Existing system c4.5 algorithm was used for performing an operation on a tree.

III. PROPOSED ALGORITHM

In our system ID3 algorithm is used in order to make an prediction about an employee performance Here we are considering several parameters in order to make an decision.

A. DATA MINING TECHNIQUES

A.1 Classification
Prediction and Estimation are types of classification techniques.

Classification Algorithms:-

Types	Names Of Algorithm.
Decision Tree	ID3, C4.5, Cart, Sprint.
Statistical	Regression, Bayesian.
Distance	Simplest Near, K Nearest Neighbors.
Neural Network	Propagation, NN supervised learning.

A.2 Clustering
Clustering is the procedure of grouping the similar elements. Below table shows different kinds of clustering algorithms which are used for grouping similar kinds of elements.

Types	Names of Algorithms
Similarity and Distance measure.	Similarity and Distance measure.
Outer	Outlier
Hierarchical	Agglomerative, Divisive
Clustering Database	Birch ,Dbscan

A.3 Association
Association rule is used inorder to find the binary variables which are frequently occurs in a database.It is used to identify the group which are related to each other. Apriori, ,CDA,DDA Algorithms are used for measurement purpose.
.

B. DATA MINING PROCESSES
Employee data are analyzed using classification method to predict the employee performance.

B.1Data Preparation
Data used here is collected from various department in order to perform an operation. Different tables are joined using a join process.

B.2Data Selection and Transformation
Fields are selected which are required for mining.Variables which are required for prediction are also collected.Here also variables collected from different databases for prediction purposes.

C. EMPLOYEE RELATED VARIABLES
A.Quality:-It refers to the quality working of an employee.This quality type information is used for prediction purpose.

B. Attendance – The extent, to which an employee is punctual, observes prescribed work break/meal periods and has an acceptable overall attendance record. Employee's willingness to work overtime as required.

C. Initiative/ Creativity – The extent to which anemployee seeks out new assignments, proposesimproved work methods, suggests ideas to eliminate waste, finds new and better ways of doing things.

D. Overall Performance – Rate employee's overall performance in comparison to position duties andresponsibilities.

D. MINING MODELS

Techniques like Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbour method etc., are used for knowledge discovery from databases. One of the popular technique among them is an classification techniques which is used for making an decision by considering an some parameters.

E.USEFULNESS OF DECISION TREE ON EMPLOYEE PERFORMANCE

A decision support tool like a decision tree that uses a tree or graph or model of decisions and their possible circumstances, By consisting an chance of event outcomes, resource costs, and utility. It is one type to display an algorithm.

Decision tree contains following points:-

1. Select a variable of training samples from training data set as nodes, branch is created to every possible value of the variables. Accordingly, the training sample set is divided into several sub-set.

2. Do the same procedure for reaming branches in order to split or divide the data into small sub-tree like structure.

It contain:-

• A leaf node-Value of the target attribute is indicated by leaf node.

• A decision node- specifies some test to be carried out on a single attribute- value, with one branch and sub-tree for each possible outcome of the test.

Information gain is used in order to selectan attribute. Information gain is calculated for each attribute in to create tree, After this select the attribute with the highest gain.

Consider two classes, P and N.

o Let the set of examples S contain p elements ofclass P and n elements of class N.

o The amount of information, needed to decide ifan arbitrary example in S belongs to P or N isdefined as

$$I(p, n) = -(p / p + n)*\log2 (p / p + n) - (n/ p +n)*\log2 (n / p + n)$$

Assume that using attribute A, a set S will be partitioned into sets {S1, S2, ..., Sn}.If Si contains pi examples of P and niexamples of N, the entropy, or the expected information neededto classify objects in all sub-trees Si is v

$$E (A) = \sum_{i=1}^{V} ((pi + ni)/(p + n))* I(pi , ni)$$

Entropy: Entropy is anexpected amount of information needed to assign a class to randomly drawn objects in S under the optimal, shortest-length code.

Information gain is calculated in order to make an reduction in entropy.

$$Gain (A) = I (p, n) – E (A)$$

IV. SIMULATION AND RESULT

In our proposed system all the task are performed by an admin. Task means adding training data set which required for calculating an entropy and information gain. Once this two module are calculated the decision tree is created by using an ID3 algorithm. After this employee data is added to the database by considering the attributes like quality, behaviour, skills, projects etc. Here we assigning a number between 1 to 10 to each attribute of an employee for calculating an performance of an employee by using an decision tree By considering the information provided by admin the probability of an employee performance is generated.
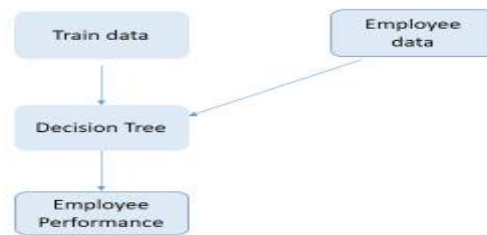
Fig- Employee performance simulation

V. CONCLUSION

This type of research provides an benefit to the organization supervisor to predict the employee's performance by considering some parameters. A current performance evaluation is required to support recommendations for merit salary adjustments and in-grade or grade change salary increases. This type also help employee who really done an magnificent work. This also helps the supervisor to find and increase the employee's performance and those employees needed special attention for reducing falling ratio for taking action at right time. Decision Tree method is used on a Employee database in order to analyse an employee data to make an prediction. Data Mining is an tool allows us to manage data in superior way. By using this system performance and quality of an management system is improved.

VI. ACKNOWLEDGEMENT

REFERENCES

[1]RohitKotalwar,RajChavan,VinaykParmwar,Sunny Gandhi," Evaluating Performance of Employee's using Classification Algorithm Based on Decision Tree." *Vol.4, No. 2, April 2014*

[2] Chen Xiaofan, Wang Fengbin," Application of Data Mining on Enterprise Human Resource Performance Management" [J]. Nanchang Hangkong University,2010(3).

[3] Dr Lakshmi Rajamani, MohdMahmoodAli[J],"Automation of decision making process for selection of talented manpower considering risk factor: A Data Mining Approach". University College of Engineering, Osmania University, Hyderabad.

[4] Han Jing, "Application of Fuzzy Data Mining Algorithm in Performance Evaluation of Human Resource". Economic and Management Institute, YanTai University, yantai, Shandong, P. R. China, 264005.

[5] E.Gothai, Dr.P.Balasubramanie,"Performance Evaluation of Hierarchical Clustering Algorithms". Department of CTPG, Kongu Engineering College, Perudurai, Tamilnadu, India.

[6] UjjwalMaulik, "Performance Evaluation of Some Clustering Algorithms and Validity Indices". Member, IEEE, SanghamitraBandyopadhyay, Member, IEEE

[7] Suo Qi. Database Marketing Research Based on DataMining[J]. Technology and Market,2007(2):53-54

[8] Li Heshan. The Investigation of the Establishment of
Enterprise Performance Management System[J].ZhongzhouCoal,2007(1):101-102

[9] Dong Yongfeng, HouXiangdan, GuJunhua, Liu Hongpu. Application of Fuzzy Data Mining in Staff Performance Appraisal[J]. Hebei University of Technology,2005(34)

[10] Ding Zhibin, Yuan Fang, DongHewei. Application of Data Mining in Higher Student Achievement Analysis[J]. Computer Engineering and Design,2006(4):590-592

[11] Wen Guiqing, JiaJinlian. Sticking Point and Mining of Human Resources in State-owned Enterprises[J],Shangxi Coal Technology,2007(1)

[12]DataMing[EB/OL].http://www.stcsm.gov.cn/learning.

BIOGRAPHY

1. Shubham Tupe is a student pursuing Bachelor of Engineering degree in Computer Engineering Department from Smt. KashibaiNavale College of Engineering,SavitribaiPhule Pune University. He Received Diploma in Computer Engineering Degree in year 2014 from Soniya Gandhi PolytechnicShrigonda,Ahmadnagar, India.

2. Chetan Mahajan is a student pursuing Bachelor of Engineering degree in Computer Engineering Department from Smt. KashibaiNavale College of Engineering,SavitribaiPhule Pune University. He Received Diploma in Computer Engineering Degree in year 2014 from SSVPS BSD Polytechnic,Dhule, India.

3. Dnyaneshwar Uplenchwar is a student pursuing Bachelor of Engineering degree in Computer Engineering Department from Smt. KashibaiNavale College of Engineering,SavitribaiPhule Pune University. He Received Diploma in Computer Engineering Degree in year 2014 from Government Polytechnic, Yavtmal, India.

4. Pratik Deo is a student pursuing Bachelor of Engineering degree in Computer Engineering Department from Smt. KashibaiNavale College of Engineering, SavitribaiPhule Pune University. He Received Diploma in Computer Engineering Degree in year 2014 from Government polytechnic,DhuleIndia.