

Intelligent Enhancement of Organization Work Flow and Work Scheduling Using Machine Learning Approach Tree Algorithm

Tareq Abed Mohammed^{†, †††}, Yaser Issam Hamodi^{††}, Naeem Th. Yousir^{†††}

[†]Collage of Engineering, Altinbas University, Istanbul, Turkey

^{††}Department of Computer Engineering, Ministry of Higher Education & Scientific Research Iraq, Baghdad, Iraq

^{†††}College of Information Engineering, Al-Nahrain University, Baghdad, Iraq

^{††††}Collage of Science, Kirkuk University, Kirkuk, Iraq

Summary

Decision Tree is one of the most used machine learning language used for analyzing business problems mostly applied to predict and decide the best way to the problems. Mostly it works on existing data and avoid repeating which is done in the past. Finding the best solution for a set of problem it will get the data of the user performance both with worst and best-case scenario. The purpose of this paper is to find the Employee Performance issues related to Educations, health, Government and many other organizations and give the solution to improve their performance and avoid hiring irrelevant employees using Decision Tree Algorithm. The type of information generated from datasheets and decided with the data processing method. More of the data which contains valuable information can be produced, Industry sector includes important information. The employee performance is calculated to find out. Well, the best way is using a favor Management and processing Staff database.

Key word:

Data Mining, Machin Learning, Intelligent System, Evolutionary Algorithm, Data Processing, Decision Tree Algorithm.

1. Introduction

Data mining is the important knowledge collection or data from databases. From this existing data different relationship among the attributes can be found. Using Data mining techniques and statistics the knowledge extract to find out the important information [1]. Employee Performance is one of the most important factors for every organization to

achieve their goals. In this report after selecting relevant attributes about Employees Performance random data are collected. Using Decision Tree Algorithm based on classification method one of the most used techniques in data mining will process the employee data and attributes to predict the Employees performance from their previous work.

Sometimes data is extracted Disclosure of Information in Databases. We can also find existing relationships and patterns. Combines data extraction Automated learning, statistics and visualization Techniques for detection and removal information. The employee has a retention. It's an industrial show Performance Management and Registry.

This project about Employees performance using Decision Tree algorithm will help the supervisor by knowing well about their Employees strength and weak points by getting this useful information one can easily assign the jobs to right person [2].

2. Method

The technique used for Employees performance is classification because this is predictive task for the system. The Decision tree algorithm based on classification method is applied for employee performance which is the most used data mine technique [5]. In this method we will collect the raw data and then transform the data in to tabular after that decision tree algorithm is applied which reduce the data and most relevant data is showed in tree [3].

This method shows the employee performance and predict task for them according to their database previous record which has a very important relation with employee performance. This method only finds the best option and will show the result using their record from which we can easily evaluate the employee from decision tree.

3. Tree Algorithm for Performance Improvement

3.1 Attributes selection Process:

Attendance: This variable is about the employee timing and to know how much punctual and serious the employee is in his job

Task: Previous task given to the employee and from this we variable we will predict his expertise in his specific fields.

Job Experience: The job experience attribute will look at his previous jobs and skill in the last organization his performance. The extent to which the employee work with other employees.

Work Quality: Quality of work for the employee that how much clearance the Employee got in the organization. The level of Employee works a professional or common. Overall Performance: Average rating for the Employee performance in the entire task [1]. This is a very important variable for calculating employee performance because sometimes it's difficult to differentiate Employee skill in specific areas [6].

Talbe 1: Work quality attribute of employee

Attendance	p	n	I (p, n)
perfect	6	2	0.356
Ordinary	3	3	1

In this table a sample data for Employee performances is given containing five different attributes. The Employee result has two values Need improvements and satisfied [4]. The result for Need improvements is 5 and for the satisfied the value is 9. In this case let suppose we gave two variable p=9 (satisfied) and n=4(Need Improvements). For this data the Information gain is calculated by the formula [11].

Talbe 2: Attendance attribute of employee

Work Quality	p	n	I (p, n)
Good	7	0	0
Bad	2	5	0.862

In table total result for Satisfied with attendance is 6,2 and with not satisfied 3,3.

4. Procedure

4.1 Information Gain

To find information Gain we have the following equation

$$I(p, n) = - (p / p + n) * \log_2 (p / p + n) - (n / p + n) * \log_2 (n / p + n) \tag{1}$$

$$I(9,5) = -(9/ 14) * \log_2 (9/ 14) - (5/ 14) * \log_2 (5 / 14) = 0.941$$

From the above table 2.2 we will calculate the entropy to decide which attribute will be the root node for decision tree [7].

4.2 Entropy

After Information Gain the next step is Entropy we can find it by the equation 2

$$Entropy E(A) = \sum_{i=1, v} (P_i + n_i) / p + n \tag{2}$$

$$E(Work Quality): 7/14 * I(2,5) + 7/14 * I(7,0) = 0.431 \text{ bits}$$

$$E(Attendance): 8/14 * I(6,2) + 6/14 * I(3,3) = 0.631 \text{ bits}$$

E

$$(Task): 7/14 * I(6,1) + 7/14 * I(3,4) = 0.938 \text{ bits}$$

$$E(Job Experience): 9/14 * I(6,3) + 5/14 * I(3,2) = 0.938 \text{ bits}$$

4.3 Gain

Now calculate the Information Gain for each attribute to draw the decision tree

$$Gain = Entropy - IG \tag{3}$$

By putting the values in equation 3

$$Gain(Work Quality) = 0.941 - 0.431 = 0.51 \text{ bits (Root node)}$$

$$Gain(Attendance) = 0.941 - 0.631 = 0.31 \text{ bits}$$

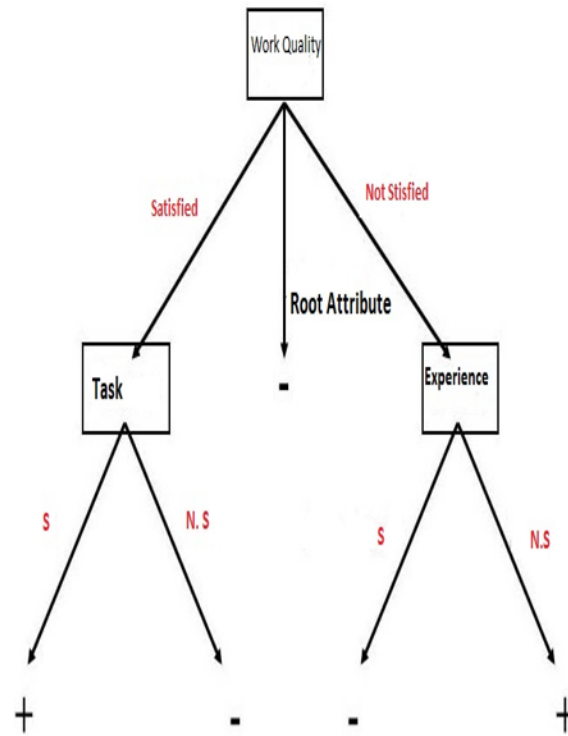
$$Gain(Task) = 0.941 - 0.938 = 0.003 \text{ bit}$$

$$Gain(Job Experience) = 0.941 - 0.939 = 0.002$$

Attendance: $I(6,2) = -(6/6+2) * \log_2 (6/6+2) - (2/6+2) * \log_2 (2/6+2) = 0.356$

$I(3,3) = -(3/3+3) * \log_2 (3/3+3) - (3/3+3) * \log_2 (3/3+3) = 1$

Now from the equation the root node is Work Quality with value 0.15 calculate the Information Gain for each attribute to draw the decision tree
For the working Quality we get the highest value in Gain it means this will be the root node of Tree Algorithm [8].



Flow Chart. 1 Attributes Values

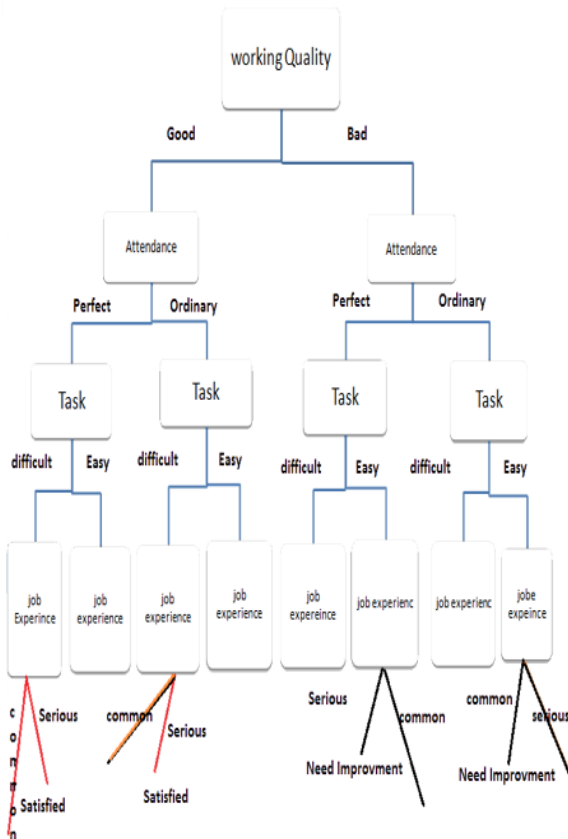


Fig. 1 Root nod Final Values

5. Result

Using decision Tree algorithm after getting the attributes and Gain values from equations finally the result is in the form of a tree contain all the attributes in sequence depending on their values [11]. In table below shows the attributes with all previous values of organizations employees which is very important for the final result which we concluded in figure 1.

The Decision Tree clearly shows that which attributes of the employees play an important role in the organization and in which attributes their low skills needs to be improvement [9]. This result will also help the company that which Employee is good for a specific task.

In figure 1 the best path for work flow starts from root node work Quality and ends to Child nodes. Each path has a different result it's clear that the red line with satisfied shows the best case for work flow improvement and need improvement is the options which should be ovoid to improve the work flow in Organizations. In table 1 the datasets for Employee is shown with their work results [10].

Table 3: Datasets of Org

No	Work Quality	Task	Experience	Result
1	Good	Easy	Serious	satisfied
2	Good	Difficult	Common	satisfied
3	Bad	Easy	Serious	Not Satisfied
4	Bad	Easy	Common	Not Satisfied
5	Good	Difficult	Serious	Satisfied
6	Bad	Difficult	Common	Satisfied
7	Good	Easy	Serious	Satisfied
8	Good	Difficult	Common	Satisfied
9	Bad	Easy	Serious	Not Satisfied
10	Good	Difficult	Serious	Satisfied
11	Bad	Difficult	Common	Not Satisfied
12	Bad	Easy	Common	Satisfied

6. Conclusion

Using this search, Ability to predict employee performance. One Evaluate the actual performance required Support for merit payments Changes in class, grade or salary Increase. At the same time, Employee and employee performance Special attention is required to reduce the rate of return to work on time. The Employees whom working quality is good and their task are difficult their performance is satisfied and the Employees whom working Quality is Bad and their task are easy in the figure 1 it's clear that their performance is poor, and they need improvements. From these attributes Employees performance can easily be differentiated, and it will improve the performance of organization by fixing these kinds of issues related to the Employees. The attributes in these projects are the basic attributes for many organizations.

References

- [1] C.F. Juang A hybrid of genetic algorithm and particle swarm optimization for recurrent network design IEEE Trans Syst Man Cybernet – Part B Cybernet, 34 (2) (2004), pp. 997-1006.
- [2] R. Ahmed, G. Karypis, "Algorithms for Mining the Evolution of Conserved Relational States in Dynamic Networks", Knowledge and Information Systems, vol. 33, no. 3, pp. 603-630, Dec. 2012.
- [3] Soryani, M and Rafat, N, 'Application of Genetic Algorithms to Feature Subset Selection in a Farsi OCR', Proceedings of World Academy of Science, Engineering and Technology Volume 18 December 2006 ISSN 1307-6884.
- [4] Mierle K, Laven K, Roweis S, Wilson G, Mining Student CVS Repositories for Performance Indicators, 2005.
- [5] Delavari N, Beikzadeh M. R, Amnuaisuk S. Application of Enhanced Analysis Model for Data Mining Processes in Higher Educational System. 6th Annual International Conference: ITEHT July 7-9, 2005, Juan Dolio, Dominican Republic.
- [6] Kalles D., Pierrakeas C., Analyzing student performance in distance learning with genetic algorithms and decision trees, Hellenic Open University, Patras, Greece, 2004.

- [7] Witten I. Frank E. WEKA Machine Learning Algorithms in Java, Morgan Kaufmann Publishers, 2000.
- [8] Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, Wirth R. CRISP-DM 1.0: Step-by-step data mining guide, 2000.
- [9] Tan P., Steinbach M., Kumar V. Introduction to DATA MINING. Pearson Education, 2006.
- [10] Ding Zhibin, Yuan Fang, Dong Hewei. Application of Data Mining in Higher Student Achievement Analysis[J]. Computer Engineering and Design,2006(4):590-592
- [11] Wen Guiqing, Jia Jinlian. Sticking Point and Mining of Human Resources in State-owned Enterprises[J], Shangxi Coal Technology,2007(1).



Tareq Abed Mohammed received B.Sc., degrees in Computer Science - College of Science, Kirkuk University, Kirkuk, Iraq, in 2007, and M.Sc degree from Cankaya University –Ankara – Turkey in 2012. In 2014 I start to Ph.D. degree in Electronic and Computer Engineering at Altinbas University. Istanbul –Turkey. Since 2015- until now, he is teaching at graduate school of science and engineering GSSE at Altinbas University.



Yaser Issam Hamodi received B.Sc., degrees in Computer Engineering - College of Electrical and Electronic Techniques, Foundation of Technical Education, Baghdad, Iraq, in 2004, and M.Sc. degree in computer engineering from Cankaya University – Ankara –Turkey in 2012.



Naeem Thjeel Yousir received B.Sc., degree in Computer Science and Statistics, Al- Rafidain University College, Baghdad, Iraq in 1998, and M.Sc degree Computer Science in 2003, and Ph.D. degree in Computer Engineering and Communications from Hacettepe University – Ankara –Turkey in 2014.