Appraisal Management System using Data mining Classification Technique

Nikhil N. Salvithal
Walchand Institute of Technology, Solapur(India)

R.B. Kulkarni, PhD
Walchand Institute of Technology,Solapur(India)

ABSTRACT
Performance appraisal may be a method within which staff are rated on however well they need met performance expectations. Performance appraisals are meant to support employees' selections, as well as promotions, terminations, training, and remuneration will increase. This task involves lots of social control selections that depends on numerous factors like human expertise, knowledge, preference and judgment. of these factors will cause inconsistent, inaccurate, difference and unpredictable selections. DM is rising knowledge analysis tool and wide utilized in order to supply valuable information for deciding. DM classification methodology are often used for talent statement, particularly for employee's job promotion. within the planned system the assorted classifier algorithms are going to be applied on Talent dataset to spot the talent set so as to judge the performance of the individual. Finally counting on accuracy one best suited classifier is chosen this method has been used to construct classification rules to predict the potential talent that helps in determining whether or not the individual is acceptable for promotion or not.

Keywords
Appraisal, data mining, classification, J48.

1. INTRODUCTION
The process to spot Associate in Nursing existing talent in organization is among the highest challenges two-faced by a company. Nowadays, there ar several areas, like in finance, medical, marketing, securities market, telecommunication, producing, health care and client relationship, that have custom-made data processing techniques[1-2][3-4][5-7]. However, the used of classification techniques in data processing approach doesn't attract abundant attention among folks in Human Resource (HR) field[8][9]. For talent prognostication, hour seeks a additional strategic role by turning to data processing strategies. methoding} is that the process of discovering new patterns from massive knowledge sets [10]. The goal information|of information} mining is to extract knowledge from an information set and gift it during a human-understandable structure.

Classification is among common technique in data processing particularly for information discovery and future set up. Classification may be a data processing technique supported machine learning that is employed to classify every item

during a set of information into a collection of predefined categories or teams [11]. A classification algorithmic rule is to use a coaching knowledge set to make a model specified the model will be wont to assign unclassified records in to 1 of the outlined categories. A check set is employed to see the accuracy of the model. Usually, the given dataset is split in to coaching and check sets, with coaching set wont to build the model and check set wont to validate it.

There are varied classifiers that ar Associate in Nursing economical and ascendible variation of call tree classification. the choice tree model is constructed by recursively ripping the coaching dataset supported Associate in Nursing optimum criterion till all records happiness to every of the partitions bear constant category label. the choice tree algorithmic rule is among the powerful classification algorithmic rule that has its advantage specified it will manufacture a model that represent explicable rules or logic statement that's additional appropriate for analyzing categorical outcomes.

theorem classifiers ar applied math supported Bayes' theorem, they predict the likelihood that a record belongs to a specific category. a straightforward theorem classifier, known as Naïve theorem classifier is comparable in performance to call tree.

Rule-based classification algorithms generate if-then rules to perform classification. This study has aimed to use these classification algorithms to handle problems on Human Talent Acquisition . This study ensures talent information acquisition by victimization employee's performance records. Talent is taken into account because the capability of any person to create a distinction in current Associate in Nursingd future performance of an organization[12]. Besides that Talent Management will be outlined because the method to confirm leadership continuity in key positions and encourages individual advancement[12]. The match of information mining issues and talent management wants ar terribly crucial. Therefore, it's important to see the appropriate data processing techniques for talent management issues.

In this study, recommendation for promotion is taken into account because the target category in classification method.

2. RELATED WORK
Classification strategies build use of mathematical and applied math techniques like call trees, applied math, neural network and support vector machines. during this section, we have a tendency to 1st review some of the connected work on time unit Talent Acquisition mistreatment data processing techniques.

Jantan, Hamdan, Othman [13] in their study Jantan et all compares the assorted classifier algorithms supported call trees to facilitate talent management. Here they predict the talent performance by means that of data noninheritable from the present databases. The performance factors for this study were supported the academician's performance. They used employees' knowledge from one in all Malaysian higher learning establishments as coaching dataset. Finally their analysis resulted by stating C4.5 because the appropriate classifier for the dataset particularly for prediction talent.

Jantan, Hamdan, Othman [14] in their study have analysed classification rule in order that necessary attributes ar known through the amount of hits for every of the attributes from generated classification rules.

Anitha Mary Florence.T and Ms.Savithri.R [12] in their work ready talent dataset administrated preliminary classification mistreatment completely different classifiers and choosen C4.5 finally for getting rule set.

Ruskova et al.[15] presents a choice web (DSS) supported formal logic model for Human Resources Appraisal and choice (HRAS). DSS are often utilized in net basedhuman resource solutions (job-search, enterprise portal, etc.). This study presents the forming the personnel of a corporation or a project team as a Fuzzy higher cognitive {process} process. Huang et al. [16] projected the appointment of managers in enterprises through fuzzy neural network, to construct a brand new model for analysis of social control talent and consequently to develop a choice web in human resource choice.

Hamdan et al. [18] during this paper projected the time unit DSS framework and therefore the potential intelligent techniques that may be accustomed develop the IDSS application in any phases of higher cognitive process processes. They represented the potential to implement the IDSS in Human Resource Management (HRM) activities. Jantan et al. [19] in their work propose resolution for a

few talent management issues mistreatment data processing techniques. They known potential talent of workers by predicting their performance mistreatment past expertise information. during this study the main target is on characteristic the patterns that relate to the talent and finding the potential mining technique for talent prediction. Their

results show that call tree or neural networks ar best suited to talent prediction.

Jantan et al. [20] suggests the potential time unit system design supported the information mining technique to resolve the talent management task by characteristic the patterns that relate to the talent by mistreatment prediction technique by manufacturing prediction rules to spot the talent. Jantan et al. [21] In their study resolved the matter of talent management by suggesting the classification and prediction techniques.. The accuracy of classifier relies on the share of check set samples that ar properly classified. The second part of experiment is to match the accuracy of classifier for attribute reduction. Here the try is to review the accuracy of the classifier for all necessary attributes. The results for full attributes show the very best accuracy for C4.5(95.1﹪)

3. PROPOSED WORK

This new system will employ classification algorithms to help decision making on appraisal & identifying key skill sets. Following steps are proposed.

1. Prepare Talent dataset.

2. Carry out classification over sample (training) data using different classifiers in acclaimed data mining tool WEKA and select most suitable.

3. Using selected classifier generate classification rule.

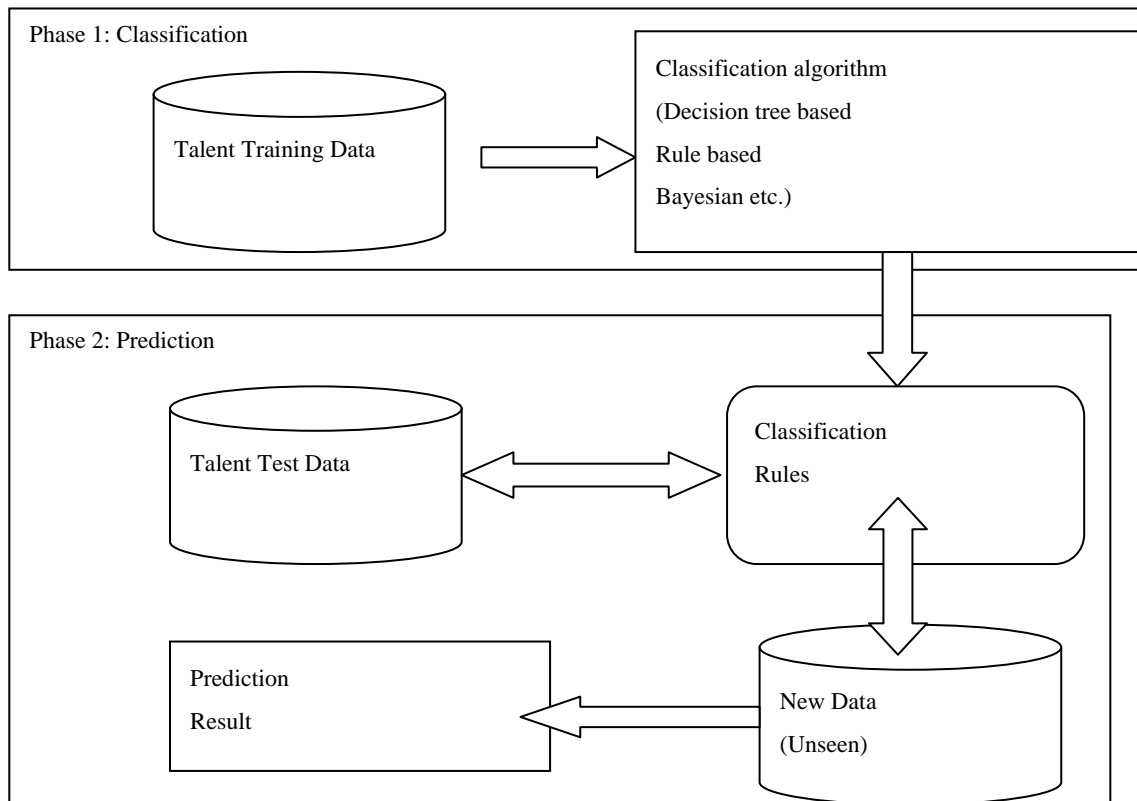4. Using generated rules predict promotionrecommendation and analyze the rules to identify key skills.

Figure 1: Experimental Setup

4. CASE STUDY USING DATA MINING CLASSIFICATION TECHNIQUE

In this case study, we focus our discussion on talent knowledge transformation component using DM classification method. Classification and prediction in DM are two forms of data analysis that can be used to extract models to describe important data classes or to predict future data trends. The classification process has two phases; the first phase is learning process whereby training data are analyzed by classification algorithm. The second phase is classification, and test data are used to estimate the accuracy of classification rules. If an accuracy is considered acceptable, the rules can be applied to the classification of new data. In this study, the experiment set up for this classification experiment is shows in fig 1.

The attributes for training dataset are selected based on the related factors for employee performance In this study, the attributes are extracted from standard performance evaluation criteria used by Pune (India) based service sector.

Table 1: Talent dataset attributes:

Attribute	Description
ID	Unique Identifier for employee
Category	P- Professional, S- Support Staff
Gender	Male and Female
Qualification	Doctorate, Master, Bachelor, Diploma and Certificate
Work Outcome	<=50%
Knowledge and Skill	<=25%
Individual Quality	<=20%
Activities and Contribution	<=5%
Evaluation mark	<=100%
Target/Class	Recommendation for promotion.

Training dataset contains 300 records. Each record holds marks for selected factors and the total i.e. evaluation marks. Data mining tool WEKA is used with this training set & test dataset of 200 records. The accuracy of the classifier is based on the percentage of the test set samples that are correctly classified. In this case, the classifier with the highest accuracy is considered as the most suitable classifier for the dataset. The classification rules are be generated using the proposed classification algorithm.

4.1 Classification result:

Appraisal Training dataset: 300 instances, Appraisal Test dataset: 200 instances.

Table 2: Evaluation of different classifiers using Training dataset on Test dataset.

Classifier	Time taken to build model	Time taken to test model	Test mode	Correctly classified instance	Incorrectly classified instances	Mean absolute Error	Root Mean squared error	Relative absolute error	Root relative squared error
Rules-OneR.	0 seconds	0.02 Seconds	User supplied Test set.	145/200 (72.5%)	55/200 (27.5%)	0.11	0.3317	58.59%	111.67%
Rules- PART	0.09 seconds	0.05 Seconds	User supplied Test set	200/200 (100%)	0/200 (0%)	0	0	0%	0%
Lazy- IBk	0 seconds	0.03 Seconds	User supplied Test set	147/200 (73.5%)	53/200 (26.5%)	0.1095	0.323	58.33%	108.74%
Bayes-NaiveBayes	0.02 seconds	0.03 Seconds	User supplied Test set	190/200 (95%)	10/200 (5%)	0.0269	0.1234	14.33%	41.53%
Functions-SMO	1.34 seconds	0.03 seconds	User supplied Test set	186/200 (93%)	14/200 (7%)	0.2444	0.3228	130.19%	108.69%

Functions-MLP	1093 seconds	0.52 Seconds	User supplied Test set	145/200 (72.5%)	55/200 (27.5%)	0.1273	0.3098	67.80%	104.32%
Trees-DecisionStump	0 seconds	0.02 Seconds	User supplied Test set	180/200 (90%)	20/200 (10%)	0.061	0.1703	32.4955%	57.3356%
Trees- J48	0.02 seconds	0.03 Seconds	User supplied Test set	200/200 (100%)	0/200 (0%)	0	0	0%	0%

From above table it can be observed that classifier J48 from decision tree family is most accurate can be easily used for prediction of recommendation for promotion on new unseen data.

J48 pruned tree Sample rules -

Work Outcome <= 34: Not eligible (206.0)

Work Outcome > 34

| Evaluation mark <= 77

| | Knowledge and Skill <= 14: Less knowledge and less efficient, cannot be considered (54.0)

| | Knowledge and Skill > 14

| | | Individual Quality <= 13: Less individual skills so no consideration for promotion (25.0)

| | | Individual Quality > 13: Knowledgeable skillful and Can improve so can be considered (3)

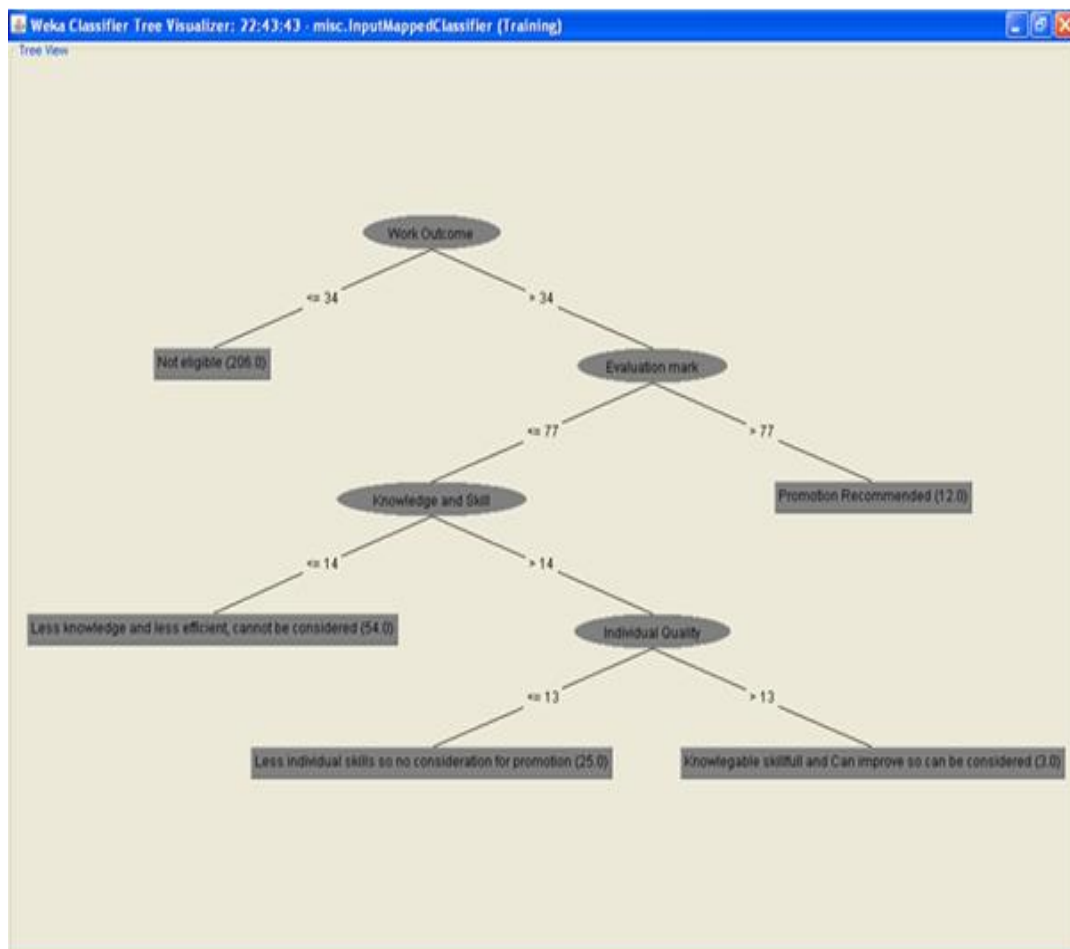| Evaluation mark > 77: Promotion Recommended (12.0)

Figure 2: Decision tree (J48)

As an example of Appraisal management using Data mining technique our application HR appraisal & analysis tool - figure 3 is based on above classification rules.
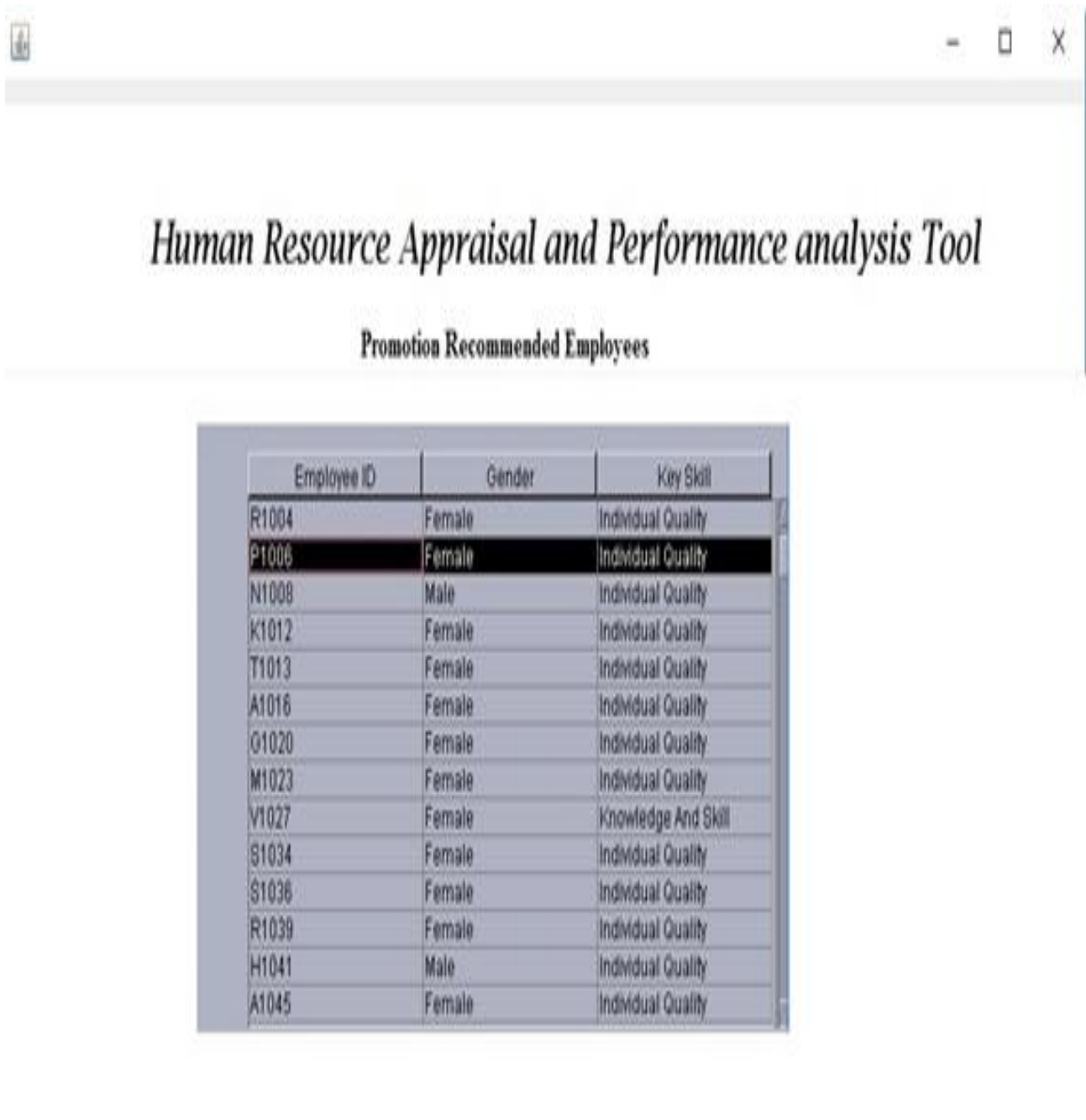

Figure 3: HR Appraisal & analysis tool.

5. CONCLUSION

With the help of data mining classification technique decision tree J48 classifier talent potentials have been predicted to help decision making on promotion and identifying important attributes. So this will help & make easy for top management to take managerial decisions & select right person for right job at right time. Also they can identify key skill attributes for appraisal process.

6. REFERENCES

[1] E. Frank, Hall, M., et al., "Data mining in bioinformatics using Weka," Bioinformatics Application Note, vol. 20, pp.2479-2481, 2004.

[2] C. Combes, Meskens, N., Rivat,, C. & Vandamme J.P., "Using a KDD process to forecast the duration of surgery,"International Journal of Production Economics, vol. 112, pp. 279-293, 2008.

[3] A. S. Chang, & Leu, S.S., "Data mining model for identifying project profitablility variables,"

International Journal of Project Management, vol. 24, pp. 199-206, 2006.

[4] S. H. Liao, Chen, Y.N., & Tseng, Y.Y., "Mining demand chain knowledge of life insurance market for new product development," Expert Systems with Applications, vol. 36, pp. 9422-9437, 2009.

[5] W. S. D. Chen, Y.K., "Using neural networks and data mining techniques for the financial distress prediction model" Expert System with Applications, vol. 36, pp. 4075-4086, 2009.

[6] C. Rygielski, J. C. Wang, and D. C. Yeh, "Data mining techniques for custome relationship management," Technology in Society, vol. 24, pp. 483-502, 2002.

[7] J. Ranjan and K. Malik, "Effective educational process: A data mining approach", VINE: The Journal of Information and Knowledge Management Systems, vol. 37, pp. 502-515, 2007.

[8] J. Ranjan, "Data Mining Techniques for better decisions in Human Resource Management Systems," International Journal of Business Information Systems, vol. 3, pp. 464-481, 2008.

[9] C. F. Chien and L. F. Chen, "Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry," Expert Systems and Applications, vol. 34, pp. 380-290, 2008

[10] SoumenChakrabarti, Earl Cox, Eibe Frank, Ralf HartmutGüting, Jaiwei Han, Xia Jiang, MichelineKamber, Sam S. Lightstone, Thomas P. NadeauRichard E. Neapolitan, Dorian Pyle, MamdouhRefaat, Markus Schneider, Toby J. Teorey, Ian H. Witten, "Data Mining-Know it all", Morgan Kaufmann Publishers, 2009.

[11] J. Han and M. Kamber , - " Data Mining; Concepts and Techniques", Morgan Kaufmann Publishers, 2000.

[12] Anitha Mary Florence.T and Ms.Savithri.R, "Talent Knowledge Acquisition Using C4.5 Classification Algorithm" International Journal of Emerging Technologies in Computational and Applied Sciences, 4(4), March-May 2013.

[13] Hamidah Jantan , Abdul Razak Hamdan and Zulaiha Ali Othman, "Intelligent DSS for Talent Management: A Proposed Architecture using Knowledge Discovery Approach", ICUIMC'12, February 20–22, 2012(ACM).

[14] Hamidah Jantan et al, "Human Talent Prediction in HRM using C4.5 Classification Algorithm" ,(IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 08, 2010.

[15] Nina At. Ruskova , "Decision Support System for Human Resources Appraisal and Selection", First International IEEE Symposium"Intelligent Systems" September 2002.

[16] Huang, L.C., et al, "Applying fuzzy neural network in human resource selection system". In Proceeding NAFIPS '04,IEEE Annual Meeting of the Fuzzy information 2004.

[17] Tai, W.S. and C.C. Hsu (2005), A Realistic Personnel Selection Tool Based on Fuzzy Data Mining Method.

[18] Jantan, H.; Hamdan, A.R.; Othman, Z.A, Potential Intelligent Techniques in Human Resource Decision Support System (HR DSS),IEEE 2008.

[19] Hamidah Jantan, Abdul Razak Hamdan,a.Zulaiha Ali Othman, Towards applyingData Mining Techniques for Talent Management, 2009International Conference on Computer Engineering and Applications.

[20] Jantan, H., Hamdan, A.R. and Othman, Z.A. (2010a), Knowledge Discovery Techniques for Talent Forecasting in Human Resource Application, International Journal of Humanities and Social Science, 5(11), pp. 694-702

[21] Jantan, H.; Hamdan, A.R.; Othman, Z.A, Classification for Talent Management Using Decision Tree Induction Techniques 2nd Conference on Data Mining and Optimization27-28 October 2009, Selangor, Malaysia.