J. Appl. Res. Ind. Eng. Vol. 7, No. 3 (2020) 267–279

Journal of Applied Research on Industrial Engineering
www.journal-aprie.com

Personnel Selection and Prediction of Organizational Positions Using Data Mining Algorithms (Case Study: Mammut Industrial Complex)

Fatemeh Mirsaeedi[1], Iman Sadeghi[2], Mohammad Ghodoosi[1,*]

[1]Department of Industrial Engineering, Faculty of Engineering, University of Torbat Heydarieh, Torbat Heydarieh, Iran.

[2]Department of Industrial Engineering, Faculty of Engineering, Iran University of Science and Technology, Tehran, Iran.

PAPER INFO

Chronicle:
Received: 09 May 2020
Reviewed: 29 May 2020
Revised: 04 July 2020
Accepted: 27 August 2020

Keywords:
Staff Selection.
Organizational Position.
Effective Features.
Data Mining.

ABSTRACT

This study aims to identify and employ qualified individuals and assign different organizational positions. Accordingly, a data mining approach is proposed. This paper presents an empirical study which has important practical application in modern human resource management. Therefore, effective features on staff selection are extracted from literature and entered into the database after expert approval respectively. Further, the impact of each feature on staff selection is determined and the ability of applied classification algorithms is compared. The results represent that the organizational position feature has a great impact on forecasting of selection or rejection. Data mining algorithms used in this study have acceptable performance based on accuracy rate, and J48 algorithm performs better comparing to other algorithms based on accuracy rate, recall, F-measure and area under Receiver Operating Characteristic (ROC) curve. Three features of background, level of education, and major are identified as effective features in association rules. Finally, an approach is presented for applying data mining algorithms in employees hiring and organizational positions assignment procedure.

1. Introduction

Today, organizations should be dynamic and flexible, because of development of potential markets, unpredictable environmental changes, increased competition, and enhanced innovation [1]. Dynamic environment arisen by modernization underlies new challenge for personnel selection which leads organizations to ignore traditional approaches and lean to modern ones. Consequently, job positions are not meant to be specified and predefined anymore. Therefore, flexible individuals are more appealing [2]. Personnel create organizational culture which is a factor that is taken into account because of its impact on corporate organizational performance [3]. So, Personnel selection plays a crucial role in human resource management. Many researchers namely [4, 5] reviewed personnel selection studies and

Mirsaeedi, F., Sadeghi, I., & Ghodoosi, M. (2020). Personnel selection and prediction of organizational positions using data mining algorithms (case study: mammut industrial complex). *Journal of applied research on industrial engineering, 7*(3), 267-279.

* Corresponding author
E-mail address: M.ghodoosi@torbath.ac.ir
10.22105/jarie.2021.233010.1170

concluded that features, including changes in enterprises, job position, employees, regulations and market, effect on staffing and recruitment [6].

Recruitment procedure can be defined as finding qualified individuals in compliance with organization requirements. Regarding to this, recruitment can be effective on growth of an organization. Recruitment is threshold of providing new workforces which can subordinate quality of human resource. Hence, staffing and selection can be considered as a competitive advantage for array of organizations from private companies to public sector whether domestic institutions or international enterprises [7].

Nowadays, there exist numerous similar companies, thus adopting efficient approaches and methods to achieve more market share is needed more than past. Suitable and qualified human resource can be a remarkable preference. Since, employee selection process depends on various features and parameters, job assignment and selection of undesirable individuals is fairly possible. Hence, determination of an ideal solution to provide human resource, recruitment and organizational position designation is vital.

In other hand, exploratory data analysis methods are being increasingly popular to handle the vast amount of data [8]. Data mining is one of the most important and applicable exploratory data analysis methods. Data mining analyze data to extract the knowledge and discover unfamiliar events, similar patterns, and relationship between data [9].

Published papers in this research area are mostly divided into, recruitment and selection, training and development [10], retention and turnover [11], performance management [12] and papers which are studied less frequent are categorized as others. This paper mainly concentrates on personnel selection. Consequently, corresponding researches are reviewed in detailed. Also, the related studies are compared with current research in *Table 1*.ome researchers, like [13] and [14] used multi-criteria decision making methods for personnel selection problem. Chien and Chen [6] used data mining for personnel selection and permanency prediction in high-tech industry. Thus, a decision tree and association rules is conducted [6]. Jantan et al. compared different classification algorithms according to accuracy rate and concluded J48 decision tree outperforms in employee's performance prediction [15]. Chen and Chien proposed an approach for personnel selection in high-tech industry which is a combination of rough set theory, support vector machine and decision tree [16]. Strohmeier and Piazza studied human resource management researches based on data mining approach. In this research, they studied papers from different dimensions including performance, methodology, data, system, user and logic [17]. Gupta and Garg rules [18] associated with applicants are weighted to prioritize job positions and a list of ranked jobs is then recommended to candidate. Sharma and Goyal proposed a decision tree and a naïve Bayes to evaluate academician's performance [19]. Sebt and Yousefi compared two approach of regression and data mining in determination of criterion effecting on personnel selection [20]. Results represent that test score, interview score, level of education, professional experience and service location are effective factors [20]. Mishra [21] used several classification algorithms in order to find well-fitted model to predict prone students for employment. They divided effective criteria into personal profile, academic information and emotional skills categories. They showed, J48 decision tree is the best suited algorithm [21]. Kirimi and Moturi [22] investigated several attributes and implemented ID3, Naïve Bayes and J48 algorithms to predict employee's performance. They found out that experience, age, gender, qualifications, training and performance score have the major impact on employee's performance and J48 decision tree performs best with the highest accuracy [22]. Kamatkar et al. [23] also studied the same problem as Kirimi and Moturi [22] and compared classification algorithms namely ID3, k-Nearest Neighborhood and J48. They showed J48 decision tree is more accurate.

Table 1. *Literature review.*

Reference	Feature										Evaluation Criteria Classification				Association Rules		
	Gender	**Age**	**Marital Status**	**Major**	**Degree**	**Experience**	**Organizational Position**	**Recruitment Channel**	**Other**	**Accuracy**	**Recall**	**F-measure**	**ROC**	**Other**	**Support**	**Confidence**	**Lift**
[6]	*	*	*	*	*	*		*								*	*
[15]	*				*				*	*							
[16]		*		*		*	*			*				*			
[18]	*	*	*	*	*	*			*	*						*	*
[19]									*	*			*				
[20]	*	*	*	*	*	*	*		*	*							
[22]	*	*			*	*	*		*	*		*	*	*			
[23]	*	*			*	*	*		*	*							
Current Article	*	*	*	*	*	*	*	*		*	*	*	*		*	*	

According to research literature, contributions of current study are as follow:

- This research studies all features integrated which have not been considered in previous papers.
- According to the literature review, accuracy rate is the most frequent evaluation criterion considered in researches. Therefore, this article considers three other evaluation criteria to distinguish the proposed algorithms precisely.
- Both criteria support and confidence are considered in order to identify the best association rules. Support criterion have not been studied in similar research.
- Due to various algorithms are different in structure and function which might lead to particular results, six algorithms are presented and the results are compared.
- A new approach to the application of data mining in the process of recruitment and assignment organizational position is presented.

The main question of this research is that, how data mining can be conducted in recruitment and organizational position assignment?

However, some other secondary questions are:

- How are the impact of each features on personnel selection?
- Do classification algorithms have acceptable accuracy to predict qualified employees for selection?
- Which algorithm outperforms the others?
- Which rules can be considered for organizational position assignment?

Regarding to this, after introduction, materials and methods are proposed, furthermore, results are discussed. Eventually conclusion remarks and further research works are presented.

2. Materials and Methods

Data mining is the process of data analysis and summarizing it as an important information in order to discover hidden patterns [24] which has received attention for converting irrelevant data to useful and acceptable information and it is also being called knowledge discovering [25]. Data mining is the tool of utilizing and exploiting useful information from a large volume of data. Due to considering data mining approach, existing hidden knowledge in datasets can be found. Data mining algorithms is categorized to clustering, classification, association rules, sequential patterns, estimation and prediction [26]. Though, this research has taken advantage of classification and association rules.

2.1. Classification Algorithms

Classification algorithms are commonly used data mining techniques and categorized as supervised learning techniques. These algorithms determine value of label variable and classify data according to results [27]. Classification algorithms apply to predict data classes (labels). These algorithms is well diversified which this research has addressed the most important ones as follow:

2.1.1. Logistic regression

It is known as statistical method to classify data according to input values. Logistic regression process begins with developing set of equations, which links input values to probability of each output classes. Placement probability of each sample in each target classes ought to be then calculated and the target class with the highest probability would be considered as predicted output. Generally, logistic regression can be applied when the variables are independent, numerical or nominal and the output variables are defined as binary [28].

2.1.2. Neural network

It is a data processing system acting like biological neural networks. Neural networks follow numerous related artificial graph edges. This method develop modelling by receiving input variables and values and changing parameters continuously [29]. Neural network is applicable for developing a prediction model which presents more accurate models despite the fact that it is time-consuming and leads to more costs. Discrete and continuous data both can be used as input values in neural networks after converting to binary format. Also, neural network can be applied to solve approximation and estimation problems which shows the results in continuous form [30].

2.1.3. K-nearest neighborhood (KNN)

A method to classify instances of a set which are located in a proximate distance and having a particular characteristics. Generally, this algorithm tries to minimize distances of data points in an appropriate class while it is maximizing distance between classes. This algorithm is quite understandable and usually takes less efforts for parameter tuning. Moreover, this algorithm is an efficient and robust approach against nuisance in dataset [31].

2.1.4. Support vector machines (SVM)

This classification method tends to calculate the maximum distance of an estimated hyperplane from data points called margin. Margin is determined by the nearest data point of each classes to the

hyperplane. Support vector machines can be implemented on both linear and non-linear solution space. Non-linear problems ought to be transformed with mapping given data into a space with higher dimension. Support vector machines are typically used for binary problems but it is also usable for problems with multiple decision values [31].

2.1.5. Bayesian

Bayesian method is a simple technique of probable classification [32]. This algorithm consider that all input criteria in dataset are independent [33].

2.1.6. J48 decision tree

Decision tree is one of the famouse classification algorithms [34]. This algorithm is a simple implementation of the C4.5 decision tree [35]. J48 develop a decision tree from training data respecting to criterion values and classify label variable [36].

2.2. Association Rules Mining

Association rules mining was introduced by [37] for the first time. Association rules mining is one of the most important data mining techniques which consist of two stages:

- Generate the frequent item sets according to minimum support.
- Extract association rules with minimum confidence [38].

2.3. Methods

This research includes several stages to implement suggested approach which is shown in *Fig. 1*. Further, each phase is defined in detailed.

2.3.1. Problem description and objective structure

Regarding to importance of personnel selection and organizational position assignment, the objective of this study is prioritizing effective features on personnel selection, assessing performance of data mining algorithms in predicting selection and evaluating association rules related to organizational position assignment.

2.3.2. Effective features extraction and data collection

Through, reviewing previous researches and inquiring expert’s viewpoint, features for predicting personnel selection are extracted. Data base can be compiled by organizational information about employees respecting to chosen features. Label variable (selection/rejection) is denoted by senior manager. Due to this, a list of staff names was given to the senior manager and satisfaction or dissatisfaction from each employee was determined. After data collection, data preparation including estimation of missing values and data normalization is used to improve performance of the algorithms.

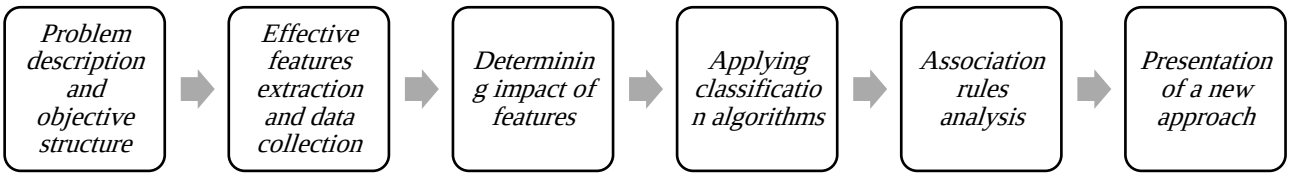

Fig. 1. *Steps of research.*

2.3.3. Determining impact of features

Various methods can be adopted to determine weight of features. Therefore, two methods of feature selection in data mining known as information gain and gain ratio are used in this research. Feature selection methods select a subset of effective features from the original features set [39]. Simplicity, independency from prediction model and interpretable output are the advantages of these algorithms [40] which are calculated by Entropy.

2.3.4. Applying classification algorithms

In order to forecast the class label, different algorithms are tested. Therefore, K-fold cross validation has been utilized. K-fold cross validation is an iterative algorithm. This algorithm separates into *K* folds which has equal number of records. This algorithm continues to run on training sets and it will terminate when all k folds are considered as a testing set [41]. An optimal number for *K* is usually suggested to be 10 when model is developed for screening and feature selection [42]. So, we dedicated a 10-fold cross validation for this problem and efficiency of algorithms is assessed by different evaluation methods. These methods include scalar and graphical methods [43]. Scalar methods such as accuracy, F-measure, and graphical methods such as area under ROC curve. Algorithm performances are assessed by accuracy rate, recall, F-measure and ROC which are following as:

– Accuracy: This is one of the criteria for evaluating classification models, which is equal to the percentage of observations that are properly categorized by the method used [44]. Accuracy is calculated as follow (*Eq. (1)*).

$$\text{Overall accuracy}=\frac{\text{TN+TP}}{\text{TP+FP+FN+TN}}. \quad (1)$$

FP is the number of the non-fault-prone instances that are misclassified as the fault-prone class. FN is the number of the fault prone instances that are misclassified as the non-fault-prone class. TP (true positive) is the number of correctly classified positive or abnormal instances. TN (true negative) is the number of correctly classified negative or normal instances. TN rate measures how well a classifier can recognize normal records. It is also called a specificity measure.

– Recall (Sensitivity): TP rate measures how well a classifier can recognize abnormal records. It is also called sensitivity measure and calculated according to *Eq. (2)*.

$$\text{Recall}=\frac{\text{TP}}{\text{FN+TP}}. \quad (2)$$

- F-measure: This criteria analyzes the effectiveness of data mining techniques and, according to *Eq. (3)*, comes from the mean harmonic between Precision (*Eq. (4)*) and Recall (*Eq. (2)*) [45].

$$\text{F-measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (3)$$

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}}. \quad (4)$$

- Area under ROC curve (AUC): The area under the ROC curve is a composite criteria indicating that the model chooses a positive position to the negative position as much as possible. The maximum is 1, and the lowest is 0.5. This criteria shows the performance of the algorithm [44].

2.3.5. Association rules analysis

As mentioned in previous section, association rules is genuinely an important approach which deals with discovering unknown pattern through dataset. In this research, we examined this approach in order to find existing correlations of organizational position assignment. Due to this, organizational position assignment is the label variable and other features are considered to be the effective features. Confidence and Support are two main criteria for evaluating rules. Confidence describes a conditional probability of variables occurrences. While, support represents a percentage of records being incurred jointly together from total frequency [46]. We considered 0.2 and 0.9 as a lower bound for support and confidence respectively which means rules with higher coefficients are recommended.

2.3.6. Presentation of a new approach for personnel selection and prediction of organizational positions

According to results, new approach for personnel selection and prediction of organizational positions is presented to use for similar researches or to be support decision tool for managers.

3. Results and Discussion

Data mining process and modelling are implemented on a collected data which is related to Mammut industrial complex staffs. Mammut industrial complex was established in 1991 tended to design and manufacture different rang of trailers, cargo, commercial vehicle applications and building systems [47]. Due to importance of human resource and recruitment of desirable individuals, Mammut Company considers employment of qualified staffs as one of the key factor in human resource management. Defined phases in methodology section are implemented on collected data from this company which are descripted below.

3.1. Effective Features Extraction and Data Collection

As mentioned former, previous papers and expert's comments can be applied in identifying effective features on selection or rejection employees. Features are represented in *Table 2*. Accordingly, staff information of Mammut company are extracted and compiled as a dataset.

***Table 2**. Details of features.*

Variable	Description	Type	Value				
Age	-	*Continuous*	*[20,70]*				
Gender	-	*Binary*	*Male*		*Female*		
Marital status	-	*Binary*	*Single*		*Married*		
Experience	*Individual's working background*	*Continuous*	*[0,20]*				
Degree	*Level of individual's education*	*Discrete*	*High School*	*Diploma*	*Associate*	*Bachelor*	*Master and Above*
Major	*Field of individual's study in educational centers and institute*	*Discrete*	*Engineering*	*Management*	*Technician*	*General*	*Other*
Recruitment channel	*Admission/ recruitment procedure*	*Binary*	*Introduced*		*Interviewed*		
Organizational position	*Individual's responsibility in organization*	*Discrete*	*Worker*	*Supervisor*	*Expert*	*Manager*	

3.2. Determining Impact of Features

Information gain and gain ratio are the two approaches used to evaluate weight of each feature. The results are shown in *Table 3*. Regarding to information of *Table 3*, organizational position is considered as the most weighted feature which indicates the importance and sensitivity of organizational position in personnel selection. On the other hand, gender is the least weighted feature reflecting that gender is not very influential on individual selection or rejection.

***Table 3.** Impact of attributes on personnel selection.*

Feature	Weight (Information gain)	Weight (Gain ratio)
Age	*0.155*	*0.288*
Gender	*0.012*	*0.016*
Marital Status	*0.054*	*0.07*
Experience	*0.191*	*0.191*
Degree	*0.134*	*0.191*
Major	*0.397*	*0.163*
Recruitment Channel	*0.024*	*0.033*
Organizational Position	*0.791*	*0.379*

3.3. Applying Classification Algorithms

Various classification algorithms are used on dataset and the results are shown in *Table 4*. Algorithms are evaluated by accuracy rate, recall, F-measure and ROC. As it is shown, all algorithms have acceptable accuracy. Logistic regression and J48 decision tree are the two algorithms with equal F-measure criteria but according to ROC, J48 decision tree and SVM have better performances. Due to better distinction, a column chart is presented as *Fig. 2*. Therefore, respecting to three evaluation criteria, J48 decision tree have the best performance.

Table 4. *Performance of classification algorithms.*

Algorithm	Accuracy	Recall	F-measure	ROC
Logistic Regression	*0.95*	*0.92*	*0.95*	*0.91*
Neural Network	*0.92*	*0.91*	*0.92*	*0.93*
KNN	*0.83*	*0.83*	*0.83*	*0.91*
SVM	*0.94*	*0.92*	*0.94*	*0.94*
Naïve Bayes	*0.91*	*0.91*	*0.91*	*0.87*
J48 Decision Tree	*0.95*	*0.93*	*0.95*	*0.94*

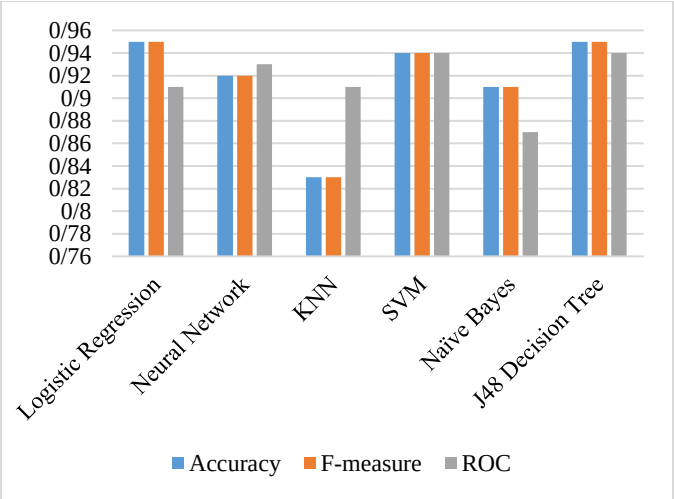

Fig. 2. *Comparison of algorithms performances.*

3.4. Association Rules Analysis

According to defined features, rules associated with organizational position assignment are extracted. Rules with the most support and confidence coefficients are listed in *Table 5*. Also, the effective features related to job positions are illustrated in *Fig. 3*.

Table 5. *Association rules for organizational position assignment.*

No.	Rule	Organizational Position	Confidence	Support
1	$IF\ 2 \leq Experience \leq 7, Degree = Bachelor$	*Expert*	*0.95*	*0.247*
2	$IF\ Experience \leq 19, Degree = High\ school\ \&\ below$	*Worker*	*1*	*0.223*
3	$IF\ Experience \geq 11, Degree = Bachelor$	*Manager*	*0.93*	*0.383*
4	$IF\ Experience \geq 15, Marital\ status = Married, Degree = Diploma$	*Supervisor*	*0.97*	*0.367*
5	$IF\ Experience \geq 8, Degree = master\ \&\ above$	*Manager*	*0.95*	*0.261*
6	$IF\ Degree = Bachelor, Recruitment\ channel = Interview, Major = Engineering$	*Expert*	*0.92*	*0.253*
7	$IF\ Degree = Master\ \&\ above, Major = Management, Experience \geq 6$	*Manager*	*0.98*	*0.246*
8	$IF\ Degree = Bachelor, Major = Engineering, Experience \geq 5$	*Supervisor*	*0.96*	*0.325*
9	$IF\ Experience \leq 5, Major = Technician$	*Worker*	*1*	*0.314*

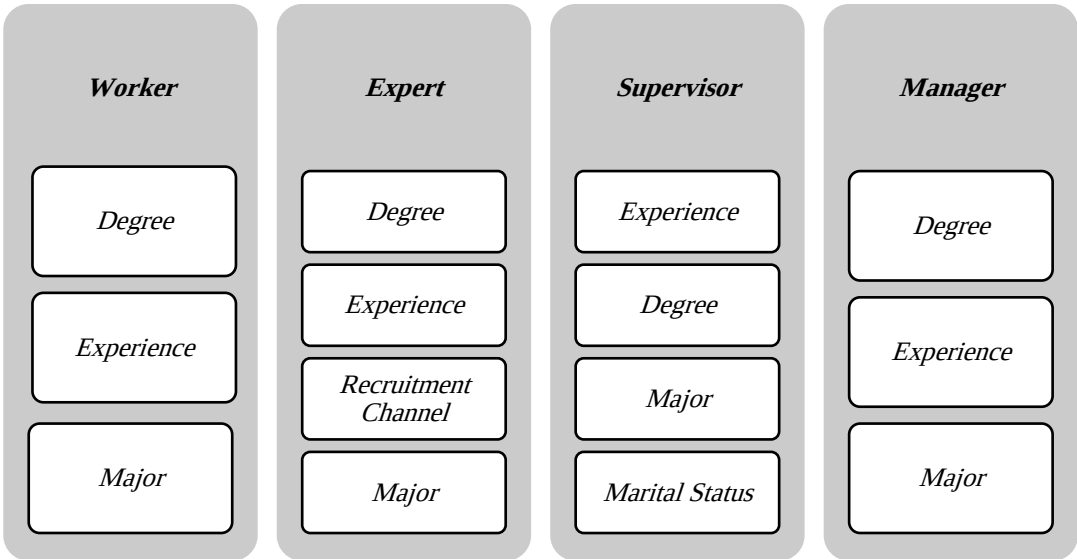

Fig. 3. *Effective features on organizational position assignment.*

However, operator job position is not addressed by any of the recognized rules. This can be emanated from insufficient volume of data related to this job position or dependency of this position to other variables such as skill which is difficult to assess. As shown in the rules, two features of experience and degree can be negated in some job positions like supervision and management. For instance, individual "A" possessing 6 years of experience with bachelor degree in engineering and individual "B" married with diploma degree and possessing 17 years of experience both are qualified for supervision. So such trade-offs can be made by decision makers. Moreover, experience, education, and major are recognized as effective features.

4. Conclusion

Importance of individual's competency in employment and job assignment is well known which inappropriate individuals or turn over can impose overwhelming costs to an enterprise. Hence, this study aims to weight effective features, evaluating performances of different data mining algorithms in predicting individual's selection or rejection and analyzing rules related to eligibility of organizational position assignment. The results represent that the organizational position weights the most among other features. This can be inferred that an individual might perform well either as an expert but as a supervisor cannot succeed that it is a logical result. Due to assessing performances of algorithms in predicting employee's selection or rejection, various data mining algorithms is considered which proposed acceptable accuracy. Also, these algorithms are evaluated by accuracy rate, recall, F-measure and ROC which J48 decision tree is introduced as the best. Respecting to organizational positions, rules are extracted and impact of each feature on position assignment is considered. As results demonstrated, data mining algorithms can be applied as decision making supporting tools for predicting individual's selection or rejection and designating them to different organizational positions. Figure 4 illustrates the procedure of applying data mining algorithms for employee recruitment and organizational position assignment. Therefore, due to decision making related to human resource management, a mechanized and intelligent method can be used along with experts and board of director viewpoints in enterprises.

Other extensions of the current research can be carried out by considering different features such as academic grades, physical condition or disability, training courses and certificates or motivational and environmental features which might be effective on predicting employee recruitment and job assignment procedure. Also, it is possible to use other data mining methods such as clustering in order to identify similar groups with similar interests and characteristics and aid human resource managers to plan for corresponding services. Presented approach can be developed for example expert system of human resources management by data mining that include all aspects of human resources management.

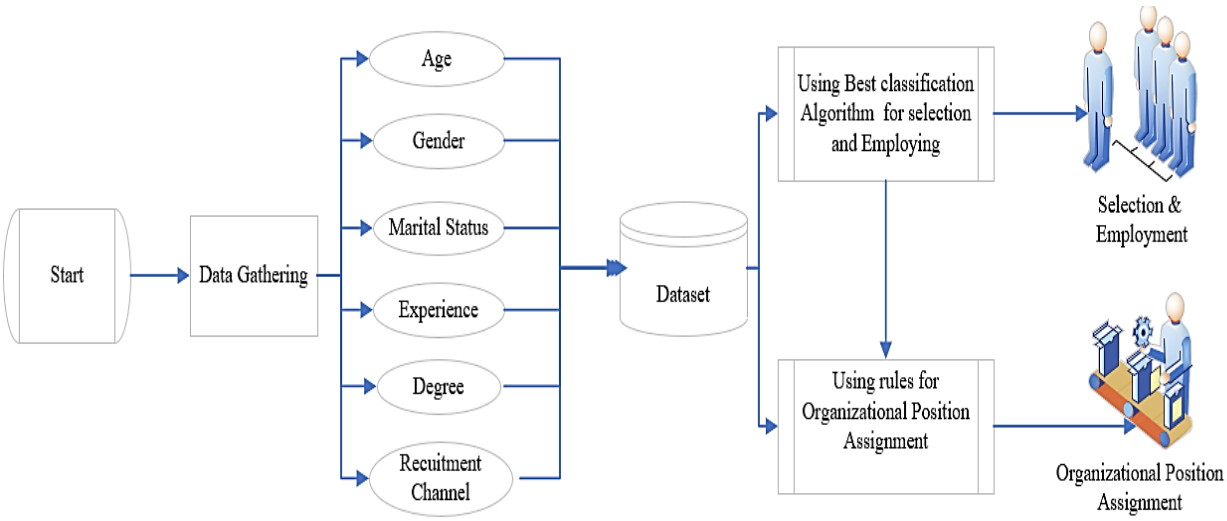

Fig. 4. *Suggested procedure.*

Refrences

[1] Fatemi, M., & Karbasian, M. (2015). Performance assessment in Isfahan municipality via knowledge management and organizational agility approach using data envelopment analysis. *Journal of applied research on industrial engineering, 2* (3), 154-167.
[2] Lievens, F. (2002). Recent trends and challenges in personnel selection. *Personnel review, 31*(5-6), 580-601.
[3] Ajripour, I., Asadpour, M., & Tabatabaie, L. (2019). A model for organization performance management applying mcdm and bsc: a case study. *Journal of applied research on industrial engineering, 6*(1), 52-70.
[4] Borman, W. C., Hanson, M. A., & Hedge, J. W. (1997). Personnel selection. *Annual review of psychology, 48*, 299-337.
[5] Robertson, I. T., & Smith, M. (2001). Personnel Selection. *Journal of occupational and organizational psychology, 74*(4), 441-472.
[6] Chien, C. F., & Chen, L. F. (2008). Data mining to improve personnel selection and enhance human capital: a case study in high-technology industry. *Expert systems with applications, 34*(1), 280-290.
[7] Santos, A. D., Armanu, A., Setiawan, M., & Rofiq, A. (2020). Effect of recruitment, selection and culture of organizations on state personnel performance. *Management science letters, 10*, 1179-1186.
[8] Shaeiri, Z., & Ghaderi, R. (2012). Modification of the fast global k-means using a fuzzy relation with application in microarray data analysis. *International journal of engineering-transactions C: aspects, 25*(4), 283-292.
[9] Darvishi, A., & Hassanpour, H. (2015). A geometric view of similarity measures in data mining. *International journal of engineering-transactions C: aspects, 28*(12), 1728-1737.
[10] Sharma, A. K., Lakhtaria, K., & Vishwakarma, S. (2013). Data mining based predictions for employees skill enhancement using pro-skill-improvement program and performance using classifier scheme algorithm. *International journal of advanced research in computer science, 4*(3), 102-107.
[11] Esmaieeli, A. M., Ghousi, R. E., & Esmaieeli, A. (2015). A data mining approach to employee turnover prediction (case study: Arak automotive parts manufacturing). *Journal of industrial and systems engineering, 8*(4), 106-121.
[12] Mohammad Naser, M., Shaaban, E., & Samir, A. (2019). A proposed model for predicting employees' performance using data mining techniques: egyptian case study. *International journal of computer science and information security, 17*(1), 31-40.
[13] Dursun, M., & Karsak, E. E. (2010). Expert systems with applications a fuzzy MCDM Approach for personnel selection. *Expert systems with applications, 37*(6), 4324-4330.
[14] Verma, M., & Rajasankar, J. (2017). A Thermodynamical approach towards group multi-criteria decision making (GMCDM) and its application to human resource Selection. *Applied soft computing journal, 52*, 232-332.
[15] Jantan, H., Hamdan, A. R., Othman, Z. A., & Puteh, M. (2010, May). Applying data mining classification techniques for employee's performance prediction. *Knowledge management 5th international conference (KMICe2010)* (pp. 645-652).
[16] Chen, L. F., & Chien, C. F. (2011). Manufacturing intelligence for class prediction and rule generation to support human capital decisions for high-tech industries. *Flexible services and manufacturing journal, 23*, 263-289.
[17] Strohmeier, S., & Piazza, F. (2013). expert systems with applications domain driven data mining in human resource management : a review of current research. *Expert systems with applications, 40*(7), 2410-2420.
[18] Gupta, A., & Garg, D. (2014). Applying data mining techniques in job recommender system for considering candidate job preferences. *International conference on advances in computing, communications and informatics (ICACCI)* (pp. 1458-1465). New Delhi, India.
[19] Sharma, M., & Goyal, A. (2015). An application of data mining to improve personnel performance evaluation in higher education sector in India. *International conference on advances in computer engineering and applications* (pp. 559-564). Ghaziabad.
[20] Sebt, M. V., & Yousefi, H. (2015). Comparing data mining approach and regression method in determining factors affecting the selection of human resources comparing data mining approach and regression method in determining factors affecting the selection of human resources. *Cumhuriyet science journal, 36*(4), 1846-1859.
[21] Mishra, T. (2016). Students ' employability prediction model through data mining. *International journal of applied engineering research, 11*(4), 2275-2282.
[22] Kirimi, J. M., & Moturi, C. A. (2016). Application of data mining classification in employee performance prediction. *International journal of computer applications, 146*(7), 28-35.
[23] Kamatkar, S. J., Tayade, A., Viloria, A., & Hernandez-Chacin, A. (2018). Application of classification technique of data mining for employee management system. *Data mining and big data* (pp. 434-444).

[24] Borkar, S., & Rajeswari, K. (2013). Predicting students academic performance using education data mining. *International journal of computer science and mobile computing, 2*(7), 273-279.
[25] Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and prediction based data mining algorithms to predict slow learners in education sector. *3rd international conference on recent trends in computing 2015(ICRTC-2015)* (pp. 500-508). Procedia Computer Science.
[26] Chang, C. L. (2007). A study of applying data mining to early intervention for developmentally-delayed children. *Expert systems with applications, 33*(2), 407–412.
[27] Mohammadi, M., Iranmanesh, S. H., Tavakoli-Moghaddam, R., & Abdollahzadeh, M. (2014). hierarchical alpha-cut fuzzy C-means, fuzzy ARTMAP and cox regression model for customer churn prediction. *International journal of engineering- TRANSACTIONS C: aspects, 27*(9), 1405-1414.
Sut, N., & Simsek, O. (2011). Comparison of regression tree data mining methods for prediction of mortality in head injury. *Expert systems with applications, 38*(12), 15534-15539.
[29] Huang, C. T., Lin, W. T., Wnag, S. T., & Wang, W. S. (2009). Planning of educational training courses by data mining: using China motor corporation as an example. *Expert systems with applications, 36*(3), 7199–7209.
[30] Rygielski, C., Wang, J. C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in society, 24*(4), 483-502.
[31] Kotsiantis, S. B. (2007). Supervised machine learning : a review of classification techniques. *Informatica, 31*, 249-268.
[32] Domingos, P., & Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine learning, 29*, 103-130.
[33] Abu Saa, A. (2016). Educational data mining & students' performance prediction. *international journal of advanced computer science and applications, 7*(5), 212-220.
[34] Movahedi Sobhani, F., & Madadi, T. (2015). Studying the suitability of different data mining methods for delay analysis in construction projects. *Journal of applied research on industrial engineering, 2*(1), 15-33.
[35] Bhuvaneswari, T., Prabaharan, S., & Subramaniyaswamy, V. (2015). An rffrctive prediction analysis using J48. *ARPN journal of engineering and applied sciences, 10*(8), 3474-3480.
[36] Panigrahi, R., & Borah, S. (2018). Rank allocation to J48 group of decision tree classifiers using binary and multiclass intrusion detection datasets. *International conference on computational intelligence and data science (ICCIDS 2018)* (pp. 323-332). The NorthCap University, India.
[37] Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD conference* (pp. 207-216). Washington, USA.
[38] Hashemzadeh, E., & Hamidi, H. (2016). Using a data mining tool and FP-growth algorithm application for extraction of the rules in two different dataset. *International journal of engineering- TRANSACTIONS C: aspects, 29*(6), 788-796.
[39] Biglari, M., Mirzaei, F., & Hassanpour, H. (2020). Feature selection for small sample sets with high dimensional data using heuristic hybrid approach. *International journal of engineering- TRANSACTIONS B: applications, 33*(2), 213-220.
[40] F Bagherzadeh, F., Ramezankhani, A., Azizi, F., Hadaegh, F., & Khalili, D. (2016). A tutorial on variable selection for clinical prediction models: Feature selection methods in data-mining could improve the results. *Journal of clinical epidemiology, 71*, 76-85.
[41] Diamantidis, N. A., Karlis, D., & Giakoumakis, E. A. (2000). Unsupervised Stratification of Cross-Validation for Accuracy Estimation. *Artificial Intelligence, 116*(1-2), 1-16.
[42] Alort, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys, 4*, 40-79.
[43] Hamidi, H., & Daraei, A. (2016). Analysis of pre-processing and post-processing methods and using data mining to diagnose heart diseases. *International journal of engineering- TRANSACTIONS A: basics, 29*(7), 921-930.
[44] Han, J., & Kamber, M. (2006). *Data mining: concepts and techniques* (2nd edition). Morgan Kaufmann.
[45] Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in human behavior, 73*, 247-256.
[46] Kotsiantis, S., & Kanellopoulos, D. (2006). Association rules mining : a recent overview basic concepts & basic association rules algorithms. *International transactions on computer science and engineering, 32*(1), 71-82.
[47] Mammutco. (2018). Mammut Industrial Group. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwiIgrCWmZLoAhWCwqYKHYjJCqYQFjAAegQIGBAC&url=http%3A%2F%2Fwww.en.mammutco.com%2F&usg=AOvVaw1llIV-ffBKUdb_JoSyBL82."