

ANALYSIS OF THE VARIABLES WHICH EFFECTS EMPLOYER'S PERFORMANCE WITH DECISION TREES AND COMPARISON OF DIFFERENT ALGORITHMS

Kürşad ZORLU

Ahi Evran University, Department of Business Administration

Abstract: This study aimed to determine the importance levels of several organizational and demographic variables influencing the perceived performances of the employees, using Classification Tree (CT) and Boosted decision Tree (BDT) methods, concerning data mining, as well as to present the usability of the tree-based algorithms in organizational studies. Decision Trees, can be evaluable as intelligent systems in data mining. When browsing the literature, it can be seen that different algorithms have been merged for decision trees. Decision Trees and tree-based algorithms are among the popular and effective methods of data mining. When looking into the literature, it may be stated that the algorithms intended for Decision Trees, generally employed in medical, engineering, and industrial researches, have recently been started to be used in social sciences, especially within the management and organization discipline. The sample of the study consisted of 10 private sector enterprises operating in Kırşehir city and having above 50 employees. According to the facts obtained from the study, Organizational Citizenship, Job Satisfaction, Life Satisfaction, and Work Environment were found to be the most important variables affecting the performances of the employees, and it was determined that the BDT algorithm presents more consistent and reliable conclusions. Thus, the study demonstrated that tree-based algorithms can be used in the correlations between the variables to form a basis for the decisions and applications of the managers.

Key Words: Data Mining, Decision Tree, Decision Making, Organizational Performance

INTRODUCTION

The process of transition from a traditional management understanding to a modern management understanding leads those efficiency-focused management types to more realistic and scientific targets. For this reason, managers have to make efficient decisions on various activities of the organization in accordance with the specified goals, under ordinary or extraordinary circum-

stances and in line with many criteria (Tohidi & Jabbari, 2012, p. 826). Changes in technologies and the possibilities provided by such changes should not be omitted for the decisions made to help achieve the goals of the organization as well as make the best choice from among the decision alternatives. Within this scope, use of knowledge/technology-intensive decision support systems may be accepted as a practical solution to make correct,

quick, and efficient decisions (Leong, 1998, p. 210). However, organizations may have limited knowledge about certain variables or may have to consider questions and environmental effects in making decisions on the organization. Some social and organizational differences that occur in employees may affect the manner in which managers make decisions. Especially, in times of crises or under heavy organizational conditions, pressure from limited time and data processing capability of an organization makes it necessary to use more dynamic and result-oriented methods in decision-making processes (Malakooti, 2011, p. 627-628; Martinson & Davison, 2007, p. 284-285; Rowell, 2009, p. 168; Wu & Pagel, 2011, p. 578-579).

When considering today's competition conditions, many managerial problems become distinct, common, or separated with many variables in various organizations operating in different sectors. It may be said that classifying many variables based on their degrees of importance in changing, regulation, improvement, and organizational development covering the entire organization to solve such problems may facilitate solving various problems and increase managerial effectiveness (Ben-David & Sterling, 2006, p. 390). During such a process, determining decision problems or variables by means of auxiliary methods, and analyzing the importance levels of such variables through traditional and quantitative methods may minimize decision-making difficulties of the managers. In particular, to raise the performance, job satisfaction, efficiency, and loyalty of the employees to the desired levels, considering such similar variables as independent variables and classifying many different variables in a sim-

pler and plain manner based on their importance levels may assist the managers in making their managerial choices correctly (Aitkenhead, 2008, p. 18-20; Malakooti, 2010; Nielsen & Nielsen, 2011, p. 185-186)

Accordingly, the methods concerning data mining ensure that the managers classify the employees with various properties according to certain variable(s), and that the togetherness of the subject-matter variables in different employee groups is analyzed. Decision Trees and tree-based algorithms are among the popular and effective methods of data mining. When looking into the literature, it may be stated that the algorithms intended for Decision Trees, generally employed in medical, engineering, and industrial researches, have recently been started to be used in social sciences, especially within the management and organization discipline. Tree-based algorithms are capable of presenting the direction and level between organizational relations in a visual environment, when compared with the statistical methods. It may be said that using the Decision Trees technique within the scope of management and organization may result in making more concrete and simpler comments in organizational researches, and may support fast and correct decision-making process. Thus, it may be possible for the managers to develop rational criteria for the candidate employees from their recruitment and formulate organizational rational strategies, policies, and tactics suitable for different employee groups (Aitkenhead, 2008, p. 19; Üngüren & Doğan, 2010, p. 42).

This study aims to classify the performance levels of the employees based on different organizational

variables and arrange such variables based on their importance levels, using classification and Regression Tree (CART) and Boosted Decision Tree (BDT) algorithms concerning Decision Trees. This study attempts to set forth the decisions regarding the usability of tree-based methods in the field of management and organization, and their advantages in organizational studies.

DATA MINING AND DECISION TREES

Data mining, which is used to solve many problems, is to select and use valuable and useful data accumulated in organizations (Özkan, 2008, p. 37). Data mining includes analyzing the data obtained from different sources using different dimensions, categorizing and summarizing such data, and determining different correlations. Such methods consist of many predictive and explanatory algorithms, such as smart system algorithms, pattern description and learning algorithms, etc. Artificial Neural Network, Genetic Algorithms, and Decision Trees are the primary methods that may be selected for data classification based on data structure and requirements (Çağiltay, 2010, p. 226-228)

Decision Trees, an effective type of data mining, is a method capable of analyzing complex and nonlinear data, and is commonly used by researchers recently (Osei-Bryson, 2007, p. 3331). With regard to enterprises, it could be noted that many studies of finance and risk management have mostly focused on uncertainty states, and those of marketing and production have been based on customer expectations (e.g., Breiman, Friedman, Olshen, & Stone, 1984; Carvalho & Freitas, 2004; Chang, Fan & Lin, 2011; Hullet,

2006; Lai, Fan, Huang & Chang, 2009; Sorensen, Miller & Oi, 2000).

One of the methods used to solve decision problems in different disciplines is Decision Trees, which are used as a problem-solving and decision-making technique, aiming to compare different alternatives based on figures and find the best method (Figure 1). Decision Trees have the ability to multi-reason, hard-to-solve, or not-fully-solved problems and determine all effective factors (İmrek, 2003, p. 323-329). When compared with other classification methods, Decision Trees have the following advantages: (a) Decision Trees produce understandable decision rules. (b) They classify without the need for extreme calculations. (c) They can be adapted to continuous and categorical variables. (d) They can be easily configured and interpreted than other similar methods (Nie et al., 2011, p. 15280). (e) There is no need to measure dependent and independent variables using the same type of criteria. The structure of a Decision Tree is similar to a flow diagram. Variables are represented by nodes. The last structures in the tree are leaves, the top structures are roots, and the branches are in between them. The basic factor determining the development of a Decision Tree is to find the variable of the root node (Lai et al., 2009, p. 3762). Decision Trees are created from top to bottom or from the general to the specific. If the data concerning the sample belong to several classes, testing is renewed at each node and division occurs. If the data belong to the same class, then the analysis terminates (Aitkenhead, 2008, p. 20; Leong, 1998, p. 213-218; Seol, Choi, Park & Park, 2007, p. 434; Yerge & Kara, 2009). Different algorithms can be cited as examples for this process. The most common Decision Tree

models are ID3 and C4.5, and the more advanced models are Classification and Regression Trees (CART) and CHAID (Chi-Square Automatic Interaction Detector) (Koyuncugil, 2007, p. 11; Seol et al., 2007, p. 434). In this study, CART has been used as a Decision Tree.

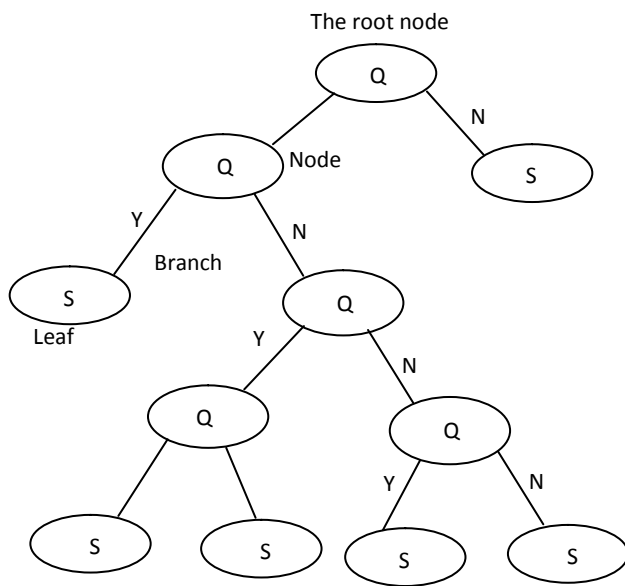


Figure 1. Representation of a Decision Tree (Aitkenhead, 2008.)

Figure 2 shows a decision-making model proposed within the scope of this study and formed by the method of Decision Trees as a basis for the strategies, plans, and applications of the managers (Can, 2005, p. 313; Chang, Fan & Lin, 2011, p. 6073; Koçel, 2007, p. 61). The manager defines the problem planned to be solved at the 1st, 2nd, and 3rd stages, and determines the target variable and other important variables over the target variable. The tree is modeled at the 4th stage by the analysis method used for a Decision Tree. The most effective variable (root node) is determined at the 5th stage. Other organizational or explanatory variables affecting the target variable in conjunction with the basic variable in the root node are arranged based on their importance levels at the 6th stage. The best decision is selected from among such formed alternatives and an application or action plan is determined at the 7th and 8th stages. If the manager fails to obtain the required results, he/she should resume the study from the 1st stage of the model.

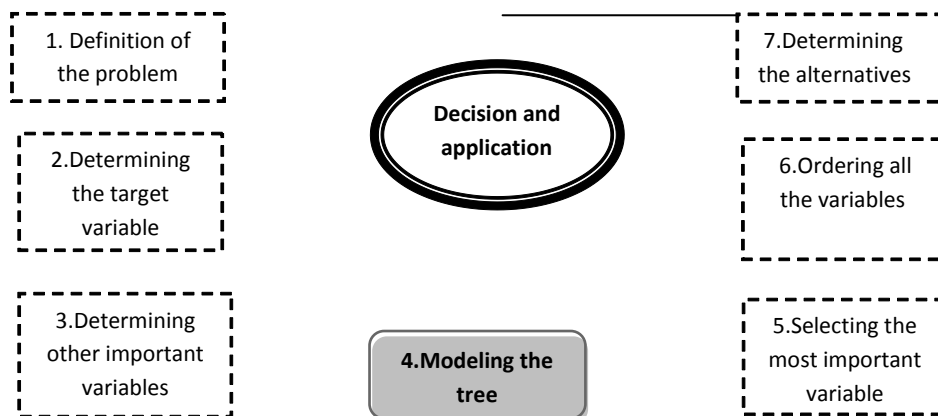


Figure 2. Decision-making process proposed for this study

CLASSIFICATION AND REGRESSION TREE

The CART algorithm, as a nonparametric method, has been used many times in medical, biological, chemical, marketing, financial studies, risk analyses, quality control, and increasingly in studies in different fields (Li, 2006, p. 262). The CART algorithm, which was first developed by Breiman et al. (1984), is capable of producing effective and understandable conclusions, especially in complex problems, on a data set, and without the need for assumptions. CART has an advantage of presenting the variables affecting the dependent variables and providing the importance of such variables in the model more easily and visually, when compared with other Classification and Regression techniques.

The CART algorithm operates in a dual structure: CT and Regression Tree. Here, what is determinant is the state of the target variable or the dependent variable. A CT is used if the target variable is categorical, and a Regression Tree is used if the target variable is continuous (Mahjoobi & Etemad-Shahidi, 2008, p. 173; Razi & Athappilly, 2005, p. 65).

CART is based on the division of each tree into two branches and forming a homogeneous tree structure with two nodes consisting of a root node (Özkan, 2008, p. 89). The CART algorithm includes three basic stages. Maximum uniformity and the best division are available at the first stage where the tree is formed (Çamdeviren et al., 2007, p. 988). Here, two sub-logarithms named Gini and Twoing are used. Gini is more sensitive to changes and aims to eliminate the variable not contributing to the model by classifying the larg-

est data set (Kurt, Ture & Kurum, 2008, p. 368). The Twoing algorithm designed for multiclass problems divides the values in the qualities into two training sets and calculates the probabilities for each of the candidate divisions. This method operates relatively slowly. Pruning occurs at the second stage of CART. Classification errors are determined by cross-verification. (Moisen, 2008, p. 584; Özkan, 2008, p. 90; p. 101-103). In this study, the Gini algorithm employed for categorical variables has been used.

TREE BOOST-BOOSTED DECISION TREES

Tree Boost is also known as “Stochastic Gradient Boosting” and “Multiple Additive Regression Trees” (MART). Mathematically, a TreeBoost model can be described as:

$$\text{PredictedTarget} = F_0 + B_1 * T_1(X) + B_2 * T_2(X) + \dots + B_M * T_M(X)$$

Where F_0 is the starting value for the series (the median target value for a regression model), X is a vector of “pseudo-residual” values remaining at this point in the series, $T_1(X)$, $T_2(X)$ are trees fitted to the pseudo-residuals and B_1 , B_2 , etc. are coefficients of the tree node predicted values that are computed by the TreeBoost algorithm (Sherrod, 2012:243).

In spite of the various advantages of Decision Trees, some algorithms known as “boosting” and combining many classifiers have been developed by Freund and Schapire to particularly increase the accuracy rate concerning the results of the study and add smart processes. Boosting is joining strong learning techniques to many numbers of simple tree models. This

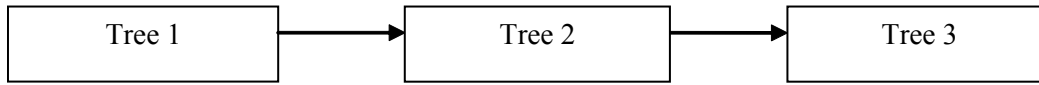


Figure 3 Tree Boosted Model (DTREG, 2012)

algorithm may increase the accuracy rate and decrease the operation time (Lemercier, Lacoste, Loum & Walter, 2012, p. 76; Neuman, Holstein, Gall & Lepage, 2004, p. 98). The Tree Boosted (TB) method, one of the tree-based methods, produces thousands of trees and applies the boosting method to regression trees (Coşkun & Karaağaoğlu, 2011, p. 185). As shown in Figure 3, every tree eliminates the error margins of the preceding tree and leads the training data to the conclusion.

It is possible to observe in various studies that Boosting algorithms give stronger and more successful results, when compared with CART and other classification methods (Hancock et al., 2005; Shin, Kim, Cho & Kang, 2012).

PERCEIVED PERFORMANCE WITH A VARIABLE TARGET AND EXPLANATORY VARIABLES

It has been observed that such methods including the opinions and comments of the persons concerning themselves, according to certain criteria, have been used more recently and named “perceived performance.” The fact that such a perception concept carries both positive and negative viewpoints together makes it possible for the employees to assess some events or variables in a different way (Carmeli, Gilat & Waldman, 2007, p. 975; Özdevcioğlu, 2003, p. 116).

Though the objectivity of the criteria developed by such studies is questionable under different

performance assessment methods, it may be accepted that they give basic clues with regard to the perfection and contribution to the job of the employees (Dewar & Kavussanu, 2011, p. 526).

In this study, performance-increasing efforts of the managers and several variables affecting performance perceptions of the employees have been arranged according to their importance levels, to contribute to their decisions. Such variables have been briefly described in the following paragraphs.

(i) Job Satisfaction

Job satisfaction is a concept implying the satisfaction and performance levels of the employees in a workplace within the frame of certain behavioral criteria, when compared with other employees (Babin & Boles, 1998, p. 82). There are many studies about the correlation between performance and job satisfaction. Most of those studies present a positive correlation between performance and job satisfaction (Arshadi, 2010; Blau, 1999; Judge, et al., 2001; Karatepe et al., 2006; Li, Natarajan & Gong, 2011; Narciss, Koerndle & Drese, 2011; Tuten & Neidermeyer, 2004; Yazıcıoğlu, 2010).

(ii) Life Satisfaction

Life satisfaction is a concept that measures the general satisfaction of a person with his/her general life, and includes a person’s physical and material well-being, interpersonal relations,

social activities, and progress (Johansson & Bernspang, 2003; Zhao, Qu & Ghiselli, 2011, p. 23). It could be observed that there are only few studies on the correlation between life satisfaction and performance than those about the correlation between job satisfaction and performance. A study conducted by Harini and Rangaiah (2010) suggests that there is no significant correlation between life satisfaction and performance levels of employees.

(iii) Work Environment

There are many studies in the literature about the effects of work environment on organizational variables and interpersonal relations. Such studies suggest that work environment affects job satisfaction, performance, absence, and similar organizational variables. It can be observed that ventilation, attention, esthetics, heat, acoustics, and lighting are mostly used with regard to work environment (Veitch vd., 2007, p. 177). This study has discussed Work Environment under two sub-variable groups: “attention and privacy area” and “esthetics and life quality,” considering the study by Veitch (2007).

(iv) Pay

Awarding and motivation are among the effective tools for orienting behaviors and providing compliance between employees and organization. The fact that making payments to employees on a flexible and changeable basis, instead of fixed wages and benefits, to decrease costs and increase motivation, can be considered as a method adopted in recent years. (Chiang & Birtch, 2010, p. 632; Eriksson & Villeval, 2008, p. 412; Yeh, Cheng & Chen, 2009, p. 1407). Studies on the correlation

between payment and performance generally suggest that Payment System and Sufficiency, as variables, have effects on employee performance (Booth & Frank, 1999, p. 448; Greenberg, 2003, p. 42-43; Widener, 2006, p. 199-204).

(v) Coach Behavior

There are only a few studies in the literature on the effect of coach behaviors on employee performance. A study conducted by Ellinger, Ellinger, and Keller (2003) suggests that positive coach behaviors have positive effects on employee performance.

(vi) Citizenship

The fact that organizational citizenship is discussed widely as a variable in recent years shows parallelism with the fact that human resources applications need to comply with competition and changes in technology. It may be said that the basic importance of organizational citizenship behaviors is its effects on employee performance, forming a basis for performance assessment system (Zhang & Jia, 2008, p. 128-130). Many studies in the literature suggest a positive correlation between organizational citizenship and the perceived performance (Chiang & Hsieh, 2012; Coleman & Borman, 2000; Podsakoff & MacKenzie, 1997; Yoon & Suh, 2003).

(vii) Empowerment

Empowerment can be considered as an effective motivation tool reflecting the level of initiative taken by the employees in an organization concerning their duties and giving behavioral results (Chow, Lo, Sha & Hong, 2006, p. 483). Empowerment not only affects the roles of employees in

an organization, but also the work environment. Trust in employees and their abilities, transfer of power, sufficient amount of flexibility, and initiative have been noted to have effects on employees' performances. Within this context, some studies suggest that there is a positive correlation between empowerment and employee performance (Chiang & Hsieh, 2012; Chow et al., 2006; Kirkman & Rosen, 1999; Linden, Wayne & Sparrowe, 2000; Niehoff, Moorman, Blakely & Fuller., 2001; Peccei & Rosenthal, 2001; Spreitzer et al., 1997).

(viii) Self-confidence

Self-confidence is a variable directing the perceptions and feelings of persons about their personalities and abilities, and having similar effects on different vocational groups (Hollenbeck & Hall, 2004, p. 257-258). There are studies suggesting that there is a positive correlation between self-confidence and performance of an employee (Bandura, 1997; Bandura & Locke, 2003; Martens, Vealey, & Burton, 1990; Vealey, 1986, 2001; Woodman & Hardy, 2003).

RESEARCH METHODOLOGY

PURPOSE AND SCOPE

This study aims to determine the demographic and organizational variables affecting employees' perceived performance levels and arrange them according to their importance, and form a basis for possible decisions of managers. The study suggests the use of CT and TB methods, which are included in tree-based algorithms, to determine the correlations between variables and utilize as a guiding analysis technique for

managers in solving various organizational/managerial problems.

SAMPLE

The sample for the study consisted of 10 private sector companies having more than 50 employees each, operating in Kırşehir province. Four of them were joint stock companies and six of them were limited liability companies. According to the information obtained from the said companies, their total number of employees was 1480 in 2011. Within this context, 750 employees were given survey forms in different days during working hours. A total of 554 survey forms were returned, out of which 436 were usable, and the rest were eliminated owing to lost data.

METHOD

The CART algorithm, which is one of the different algorithms of Decision Trees, and TB, which makes tree-based methods gain the learning ability, were employed as the analysis method in this study. The analysis was conducted using the package program DTREG 3.0, offering a fast and visual use, especially when classification and tree-based methods are preferred and developed by Sherrod.

It is possible to use all the ranges in the 5-point Likert scale to determine the effects on Performance, which was selected as the target variable in the study. However, such a case increases the number of branching in the tree more than required and causes unnecessary loss of time. Accordingly, as shown in Figure 4, the target variable value to be used in the analysis was divided into three main sections categorically. The values from 1 to 2 were determined to be "bad" performance, 2 to

4 were considered as “average” performance, and 4 to 5 were considered as “good” performance. Thus, the target variable must be categorical, for

the first time in the study, when using CT. If the target variable is continuous, instead of categorical, then Regression Tree is used.

1		2	3	4		5
	Bad		Average		Good	

Figure 4 Scale used for perceived performance level in the analysis

Within this context, GINI sub-logarithm was used, and Y represents the perceived performance and $X_1, X_2, \dots, X_{N-1}, X_N$ represent the explanatory variables. DTREG allows the same data to be used in different tree-based methods. Second, categorical and continuous data were analyzed using the BDT method so as to allow comparison with CT. Importance levels of the variables and some performance criteria were used for comparison.

Accuracy, Precision, Recall, and F-measure values were considered as performance criteria. Accuracy is the difference between the number of incorrectly classified samples and total number of samples. Precision is the difference between the number of correctly classified positive samples and total positive samples. F-measure is the harmonic mean of accuracy and precision.

As BDT includes many trees, only a tree model of CT is given in Appendix B and Appendix C. Explanatory variables of the study were examined under two sections: demographic variables and behavioral variables. Behavioral variables included Job Satisfaction, Life Satisfaction, Organizational Citizenship, Coach Behavior, Self-confidence, Empowerment, Work Environment, and Pay. Demographic variables comprised Gender, Edu-

cation, Marital Status, Age, Position, Working Manner, Service Period, and Union Membership. As the CART logarithm allows the same target variable to direct different inputs, the effects of organizational and demographic variables on the target variable were examined in two inputs. It may be said that such a separation would allow managers to make more comprehensive and balanced decisions.

VARIABLE IMPORTANCE TABLE

In the study, tables which show the importance of organizational and demographic variables are calculated for two algorithms one by one. That sequence, which shows importance level, can be respected as an advantage in the concept of DTREG.

Importance scores are computed by using information about how variables were used as primary splitters and also as surrogate splitters. Obviously, a variable that is selected as a primary splitter early in the tree is important. What is less obvious is that surrogate splitters that closely mimic the primary splitter are also important because they may be nearly as good as the primary splitter in producing the tree. If a primary splitter is slightly better than a surrogate, then the primary splitter

may “mask” the significance of the other variable. By considering surrogate splits, the importance measure calculated by DTREG gives a more accurate measure of the actual and potential value of a predictor. The importance score for the most important predictor is scaled to a value of 100.00. Other predictors will have lower scores. Only predictors with scores greater than zero are shown in the table. DTREG, takes 100.00 value for the most important predictor. Other predictors’ scores might be lower from that ratio. Only, predictors which has ratios bigger than 0 has been showed in the tables (Sherrod, 2012: 206).

SURVEY FORM AND CRITERIA

A survey was used in the study to obtain the data. Survey forms consisted of questions aiming to determine the demographic and organizational variables. Demographic variables were Gender, Education, Marital Status, Age, Position, Working Manner, Service Period, and Union Membership. When determining the organizational variables, the sources presented in Table 1 were utilized. In this study, the Likert scale consisting of 1. “I do not share at all”; 2. “I share a little”; 3. “I do not have opinion”; 4.”I share lot”; and 5.”I share completely” was employed preferred.

Table 1 Studies from which the statements included in this survey were used.

Variables	Researches
Perceived performance	Babin & Boles 1996; Karatepe et al.,2006
Organizational citizenship	Podsakoff et al.,1990
Coach behavior	Ellinger, Ellinger & Keller, 2003
Work environment	Veitch et al.,2007
Empowerment	Hartline & Ferrell 1996
Job satisfaction	Chiang & Birtch, 2010
Life satisfaction	Diener et al.,1985
Self-confidence	Beattie et al., 2011; Gabriel, Renaud & Tippin, 2007
Pay and promotion	Chiang & Birtch, 2010

VALIDITY AND RELIABILITY

Factor analysis and Cronbach’s alpha (α) coefficients were determined to establish the validity and reliability of the survey form. A factor analysis was conducted and those factor statements with a factor load below 0.40 were removed from the analysis. Accordingly, the factor analysis results of the variables and statements in the survey and

Cronbach’s alpha values are given in Appendix A. Based on both factor loads and Cronbach’s alpha values, it can be considered that the survey form is acceptable.

DEMOGRAPHIC PROPERTIES

Frequency distributions and percentage values were utilized to determine the demographic properties in this study. From Table 2, it can be observed

that 83% of the employees who participated in the survey are men with a concentrated age range of 28–37 years, 67.2% are married, 65.4% are high-school graduates, 67.8% have a service

period of less than 5 years, 45.3% are workers or office attendants, 55.5% are not members of any union, and 56.2% are full-time employees.

Table 2 Demographic Properties Regarding Both the Samples.

Factor	Frequency	%	Factor	Frequency	%
Gender			Union membership		
Male	361	83	Yes	195	44.5
Female	75	17	No	235	55.5
Age (years)			Service period		
18–27	158	36.2	Less than 1 year	97	22.3
28–37	196	45	Between 1 and 2 years	77	17.7
38–47	60	13.8	Between 2 and 5 years	121	27.8
48–57	21	4.5	Between 5 and 10 years	98	22.5
58–67	1	0.2	10 years and above	43	9.7
Position			Education		
Upper-level manager	16	3.6	Primary school	32	7.3
Middle manager	34	7.8	Middle school	72	16.5
Lower-level manager	33	7.6	High school	285	65.4
Workers or office attendants	202	45.3	Associate degree	36	8.3
Technical staff	123	28.3	Bachelor degree	8	1.8
Foreman	28	6.4	Postgraduate degree	3	1
Marital status			Working manner		
Married	293	67.2	Full-time	245	56.2
Single	128	29.4	Part-time	7	2
Widow	15	3.3	Shift	179	41
			Weekends	5	1
Total	436	100	Total	436	100

RESULTS

FINDINGS CONCERNING ORGANIZATIONAL VARIABLES

With regard to the importance levels of the organizational variables used in this study (Table 3), it

can be observed that Organizational Citizenship, Job Satisfaction, and Life Satisfaction are the most important variables in both the methods. In particular, Organizational Citizenship forms the root node for both the methods. Furthermore, in both the methods, according to the importance

level, Attention Secrecy and Pay variables are behind. However, some differences in classifying based on other variables could be noted. CT method found Empowerment to be unimportant, while BDT method observed it to be the fourth

important variable. It can be seen that Esthetics-life quality and Attention Secrecy are less important variables in both the methods. In addition, it can be noted that the weight values of the variables in the BDT method are relatively higher.

Table 3 Overall Importance of Organizational Variables

	BDT	Importance	CT	Importance
1	Citizenship	100.000	Citizenship	100.000
2	Job satisfaction	45.131	Job satisfaction	30.874
3	Life satisfaction	44.067	Life satisfaction	26.221
4	Empowerment	41.707	Esthetics-life	20.157
5	Esthetics-life	40.912	Coach behavior	20.121
6	Coach behavior	35.529	Self-confidence	18.802
7	Pay	34.513	Pay	6.792
8	Self-confidence	28.567	Attention secrecy	5.378
9	Attention secrecy	28.563	Empowerment	Unimportant

As shown in Table 4, with regard to the comparisons between the performance criteria of both the methods, it can be observed that Accuracy, Precision, Recall, and F-measure values of the findings of BDT method are higher in the three

performance classifications. Thus, it is possible to suggest that BDT analysis method is able to explain the target variable better in forms of all the criteria.

Table 4 Comparison between the performances of BDT and CT methods in organizational variables group

	Accuracy		Precision		Recall		F-measure	
	<i>BDT</i>	<i>CT</i>	<i>BDT</i>	<i>CT</i>	<i>BDT</i>	<i>CT</i>	<i>BDT</i>	<i>CT</i>
Good level	0.918	0.809	0.922	0.86	0.922	0.762	0.922	0.804
Bad level	0.985	0.949	1.00	0.68	0.713	0.502	0.829	0.577
Mid-Level	0.916	0.809	0.882	0.733	0.924	0.857	0.903	0.791

FINDINGS OBTAINED FROM DEMOGRAPHICAL VARIABLES

The effect weights of the two algorithms were also found to be different in the findings concerning demographical variables. With regard to the target variable's explanation level of demographical variables (Table 5), it can be seen that the Position is the most important variable for both the methods, and

Gender variable is unimportant in both the methods. Education is the second important variable in the BDT method, while Marital Status is the second important variable in CT method. Arrangement of other variables based on their importance levels appears to be without change. Similarly, it can be noted that Union Membership is the second unimportant variable following Gender in both the methods.

Table 5 Overall Importance of Demographic Variables

	BDT	Importance	CT	Importance
1	Position	100.000	Position	100.000
2	Education	65.811	Marital status	96.517
3	Service period	58.142	Service period	71.980
4	Working manner	48.506	Working manner	38.984
5	Age	40.614	Age	34.563
6	Marital status	32.770	Education	33.275
7	Union membership	25.312	Union membership	11.783
8	Gender	Unimportant	Gender	Unimportant

Table 6 presents a comparison of the target variable explanation regarding the power and success of demographical variables. Accordingly, with regard to the success criteria of Accuracy,

Precision, Recall, and F-measure, it can be seen that the BDT algorithm is better in all the three classifications.

Table 6 Comparison of performances of BDT and CT methods with respect to demographic variables.

	Accuracy		Precision		Recall		F-measure	
	TB	CT	TB	CT	TB	CT	TB	CT
Good level	0.749	0.610	0.698	0.582	0.894	0.843	0.784	0.688
Bad level	0.934	0.931	1.00	0.680	4.17	0.706	0.080	0.508
Mid-Level	0.750	0.623	0.749	0.5841	0.610	0.360	0.671	0.445

DISCUSSION

Actually, the algorithms which used in the study rose on three basic principles: (1) Decision tree

technique; can be used as an alternative way for specifying the relation between organizational variables and the importance level of variables.

(2) Therefore, single-tree and tree-boost algorithms would be chosen. (3) However, (in organizational studies, decision tree technique) in order to rise the safety level single tree technique and which enhances its validity, it should be used with tree-boost algorithms. Nonetheless, while single-tree can provide a visual tree; tree-boost, which is a high ratio safety level and consistent results, cannot provide a visual tree due to its over hundred operations and branches. Ultimately, tree boost algorithm, can be used in order to improve predictable accuracy like single-tree.

Citizenship, Job Satisfaction, Life Satisfaction, and Pay variables in the organizational variables group have the same order of importance as that of Position, Service Period, Work Manner, and Age in the demographical variables group. The most important organizational variables affecting the performances of the employees are Citizenship, Job Satisfaction, and Life Satisfaction. The fact that Pay variable is behind is an important finding. However, such a finding might be owing to the participants not being objective when responding to the survey. It could be noted that the occupied "Position" is the most important explanatory variable in the demographic variables group, while Gender and Union Membership are the most unimportant variables. Accordingly, it is possible to suggest that managers are required to make decisions and applications, which may increase Organizational Citizenship, Job Satisfaction, and Life Satisfaction to improve the performances of the employees, and while doing so, they should consider different positions, education levels, and service periods of the employees. In fact, it must be stated that less important variables observed in this study are capable of affecting those three

important variables, because, variables such as Pay, Coach Behavior, and Empowerment are the factors capable of affecting Job Satisfaction and Life Satisfaction variables of the employees.

CONCLUSIONS

In this study, using CT and BDT methods, some organizational and demographic variables affecting the performance levels (perceived performance) of the employees were arranged based on their importance levels, and some algorithms used were compared based on several success criteria. It is possible to reach three important conclusions based on the findings of this study: (a) Managers can use CT and BDT algorithms in fields that include many data sets, such as production, finance, and marketing, as well as in managerial and organizational fields, and especially in analyzing data from surveys. Managers can form organizational and problem-related decision alternatives using tree-based methods, and can determine many action plans. Especially, it is possible to quickly present the effect or importance of many explanatory variables (independent variables) based on a target (dependent) variable. This method may bring to light many classifications and possible correlations that are difficult to determine. Besides, it can be asserted that tree-based algorithms provide an important support for statistical methods used frequently in organizational studies. (b) The fact that using multiple algorithms in organizational studies when employing tree-based methods may ensure that more reliable results are obtained.

Furthermore, the two algorithms used in this study are able to put certain variables in different orders, thus demonstrating the importance of

this approach. When considering that there are many algorithms in data mining, it will be more realistic for researchers to act based on research methods, participants' state, and data types when selecting the most optimum algorithms. (c) It is understood that there are some differences between the findings and accuracy levels presented by CT and BDT algorithms. Within this context, with regard to the order of importance levels of explanatory variables and success criteria used in the analyses, it can be suggested that BDT algorithm presents more reliable and consistent results in correlations. This finding is consistent with that reported by similar studies in the literature. However, it is very important for future studies

to employ different organizational variables and similar tree-based algorithms.

LIMITATIONS

Use of tree-based methods in the field of management and organization is rather new. The fact that there are a very few similar studies in the literature can be accepted as a limitation. Another limitation is that the participants might have acted subjectively when responding to the survey.

APPENDICES

Appendix A Factor Loadings Concerning and Cronbach's Alpha (α) Coefficients

<i>Perceived performance ($\alpha=0.934$) Loadings</i>
Managing your working time optimally 0.808
Your work discipline 0.864
Your decision-making speed on your work 0.895
Your problem-solving skill concerning your work 0.880
Your determining and achieving a target 0.774
Your continuous contribution to your enterprise for development 0.864
Your using resources cost-effectively 0.818
Your success in team work 0.814
Your success in determining your customers' expectations 0.659
<i>Job satisfaction ($\alpha=0.857$)</i>
I find my job pleasant 0.815
My job satisfies me 0.823
I like my job 0.829
<i>Life satisfaction ($\alpha=0.797$)</i>

I have an ideal life in many respects 0.643

My life conditions are pretty satisfactory 0.847

My life conditions satisfy me 0.869

I have obtained what I want from life recently 0.772

If I could make my life from the beginning, I would almost change nothing 0.665

Work environment

Attention and privacy zone ($\alpha=0.817$)

Amount of noise arising from other peoples' dialogues when working at the desk. 0.681

Frequency of distraction by other people 0.774

Distance between me and people I work with 0.800

Privacy of talking in my work place 0.756

Capacity of my personal working area to accommodate tools-materials and guests etc. 0.653

Esthetics and Life Quality ($\alpha=0.882$)

Freedom to change physical conditions in my working area 0.615

Esthetic appearance of the work place 0.700

Air follow and its quality in my work area 0.847

Heat in my work area 0.794

Lighting quality in my work area 0.748

Access to outside view 0.770

Pay ($\alpha=0.892$)

My salary 0.769

Incentives and rewards 0.833

Bonus rate 0.854

Salary increase 0.851

Promotion possibilities 0.823

Coach behavior ($\alpha=0.851$)

My manager asks my expectations and communicates with me about the importance of such expectations for long-term targets of the company. 0.777

My manager helps me see the big picture concerning my job, and encourages me to develop my perspective. 0.872

My manager gives me feedback. 0.868

My manager provides me with resources to do my work in the best way. 0.799

Organizational citizenship ($\alpha=0.992$)

I try to do my job in the best way 0.815

I obey the rules and the applications of the company even though I don't think that nobody supervises me. 0.844

I keep the notices, information given by the company. 0.762

I am careful about how my behaviors affect other peoples' works. 0.829

I avoid having problems with my colleagues. 0.866

I always try to help my circle. 0.852

Empowerment ($\alpha=0.774$)

I don't have to deal with the problems of the management when I deal with the customer problems 0.729

I can use my own judgment when I deal with the customer problems. 0.829

I am encouraged here about using my own initiative 0.851

Here, creative thinking is encouraged about solving problems. 0.723

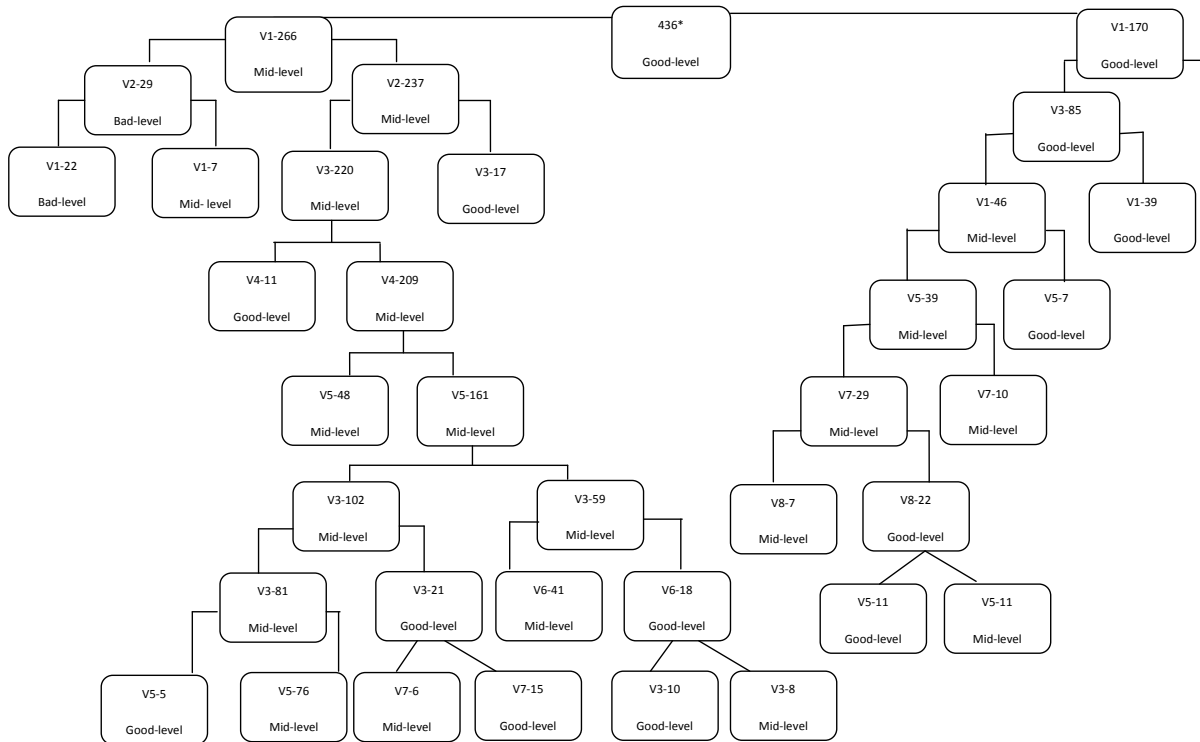
Self-confidence ($\alpha=0.662$)

My self-confidence does not change most, it generally remains same. 0.789

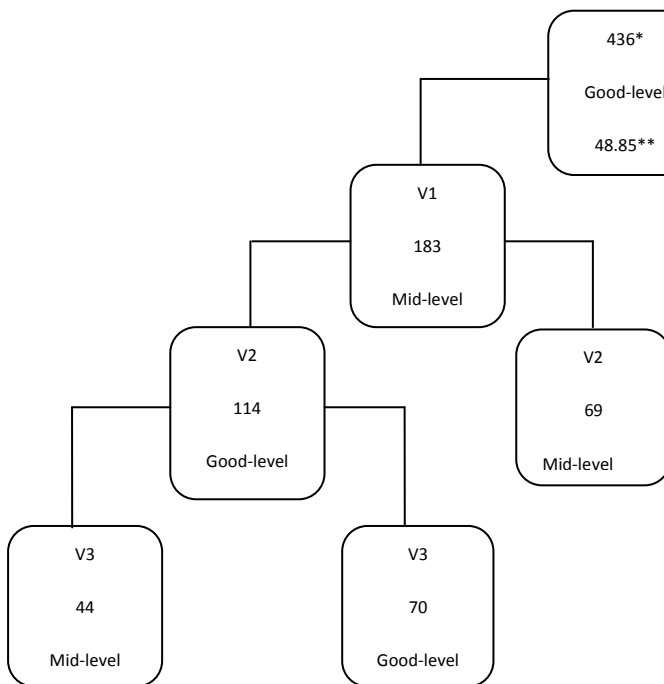
My self-confidence will not be affected by the results from competition. 0.849

I am sure about my self-confidence and skills. 0.701

Appendix B. Tree Model of Organizational Variables



Appendix C. Tree model of demographic variables



Note: *N ** Misclassification

V1: Position

V2: Marital status

V3: Education

REFERENCES

- AITKENHEAD, M. J. (2008).** A co-evolving decision tree classification method. *Expert Systems with Applications*, 34, 18-25.
- ARSHADI, N. (2010).** Basic need satisfaction, work motivation, and job performance in an industrial company in Iran. *Procedia Social and Behavioral Sciences*, 5, 1267-1272.
- BABIN, B. J., AND BOLES, J. S. (1998).** Employee behavior in a service environment: A model and test of potential differences between men and women. *J Mark*, 62, 77-91.
- BANDURA, A. (1997).** Self-efficacy: The exercise of control. New York: W. H. Freeman.
- BANDURA, A., AND LOCKE, E. A. (2003).** Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88, 87-99.
- BATEMAN, T. S., AND ORGAN, D. W. (1983).** Job satisfaction and the good soldier: The relationship between affect and employee citizenship. *Academy of Management Journal*, 26, 587-595.
- BEN-DAVID, A., AND STERLING, L. (2006).** Generating rules from examples of human multiattribute decision making should be simple. *Expert Systems with Applications*, 31, 390-396.
- BOOTH, A. L., AND FRANK J. (1999).** Earnings, productivity, and performance-related pay. *Journal of Labor Economics*, 17, 447-463.
- BREIMAN, L., FRIEDMAN, J., OLSHEN, R., AND STONE, C. (1984).** Classification and regression trees. Belmont, CA: Wadsworth.
- CAN, H. (2005).** Organizasyon ve Yönetim. Ankara: Siyasal Kitabevi.
- ÇAMDEVİREN, A. H., YAZICI, A. C., AKKUŞ, Z., BUĞDAYCI, R., AND SUNGUR, M. A. (2007).** Comparison of logistic regression model and classification tree: An application to postpartum depression data. *Expert Systems with Applications*, 32, 987-994.
- CARMELI, A., GILAT, G., AND WALDMAN, D. A. (2007).** The role of perceived organizational performance in organizational identification, adjustment and job performance. *The Journal of Management Studies*, 44, 972-999.
- CARVALHO, D. R., AND FREITAS, A. A. (2004).** A hybrid decision tree/genetic algorithm method for data mining. *Information Sciences*, 163, 13-35.
- ÇELİK, C., AND KARADAL, H. (2007).** Kobi'lerin Sorunları ve Çözüm Stratejilerinin Algılanan Performans Üzerine Etkileri: (Aksaray Ve Mersin Örneği), Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi, 16(2), 119-138.
- CHANG, P., FAN, C., AND LIN, J. L. (2011).** Trend discovery in financial time series data using a case based fuzzy decision tree. *Expert Systems with Applications*, 38, 6070-6080.
- CHIANG, F. T., AND BIRTCH, T. A. (2010).** Pay for performance and work attitudes: The mediating role of employee-organization service value congruence. *International Journal of Hospitality Management*, 29, 632-640.

CHIANG, C. F., AND HSIEH, T. S. (2012).

The impacts of perceived organizational support and psychological empowerment on job performance: The mediating effects of organizational citizenship behavior. *International Journal of Hospitality Management*, 31, 180-190.

CHOW, I. H., LO, T. W., SHA, Z., AND HONG, J. (2006).

The impact of developmental experience, empowerment, and organizational support on catering service staff performance. *Hospitality Management*, 25, 478-495.

COLEMAN, V. I., AND BORMAN, W. C. (2000).

Investigating the underlying structure of the citizenship performance domain. *Human Resource Management Review*, 10, 25-44.

COŞKUN, E., AND KARAAĞAOĞLU, E. (2011).

Microarray gene expression data analysis with data mining methods. *Hacettepe Tıp Dergisi*, 42, 180-189.

ÇAĞILTAY, N. E. (2010). İş Zekası ve Ver Ambarı Sistemleri. Ankara: ODTÜ Yayıncılık.

DEWAR, A., AND KAVUSSANU, M. (2011).

Achievement goals and emotions in golf: The mediating and moderating role of perceived performance. *Psychology of Sport and Exercise*, 12, 525-532.

DTREG, <<http://www.dtrege.com/svm.htm>>. Last accessed April 2012.

ELLINGER, A. D., ELLINGER, A. E., AND

KELLER, S. B. (2003). Supervisory coaching behavior, employee satisfaction, and warehouse employee performance: A dyadic perspective

in the distribution industry. *Human Resource Development Quarterly*, 14(4), 435-458.

ERIKSSON, T., AND VILLEVAL, M. C. (2008).

Other-regarding preferences and performance pay. An experiment on incentives and sorting. *Journal of Economic Behavior and Organization*, 68, 412-421.

GREENBERG, J. (2003). Creating unfairness

by mandating fair procedures: The hidden words of a pay-for-performance plan. *Human Resource Management Review*, 13, 41-57. [http://dx.doi.org/10.1016/S1053-4822\(02\)00098-0](http://dx.doi.org/10.1016/S1053-4822(02)00098-0)

GÜRBÜZ, S., AND YÜKSEL, M. (2008). Çalışma

Ortamında Duygusal Zeka: İş Performansı, İş Tatmini, Örgütsel Vatandaşlık Davranışı ve Bazı Demografik Özelliklerle İlişkisi, *Doğuş Üniversitesi Dergisi*, 9, 174-190.

HANCOCK, T., PUT, R., COOMANS, D., VANDERHEYDEN, Y., AND EVERINGHAM, Y. (2005).

A performance comparison of modern statistical techniques for molecular descriptor selection and retention prediction in chromatographic QSRR studies. *Chemometrics and Intelligent Laboratory Systems*, 76, 185-196.

HARINI, C., AND RANGAIAH, B. (2010).

The influence of emotional intelligence on life satisfaction and job performance of employees. *Asian Journal of Development Matters*, 4(2), 177-183.

HOLLENBECK, G., AND HALL, D. T. (2004).

Self-confidence and leader performance. *Organizational Dynamics*, 33(3), 254-269.

- HULLET, D., (2006).** “Decision trees analysis for the risk averse organization”, available on-line at http://www.projectrisk.com/Welcome/White_Papers
- İMREK, M. K. (2003).** Yöneticiler İçin Karar Verme Teknikleri. İstanbul: Beta Yayınları.
- JUDGE, T. A., THORESEN, C. J., BONO, J. E., & PATTON, G. K. (2001).** The job satisfaction–job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, 127, 376-407.
- KARATEPE, O. M., ULUDAG, O., MENEVIS I, HADZIMEHMEDAGIC, L., & BADDAR, L. (2006).** The effects of selected individual characteristics on frontline employee performance and job satisfaction. *Tourism Management*, 27, 547-560.
- KIRKMAN, B. L., & ROSEN, B. (1999).** Beyond self-management: Antecedents and consequences of team empowerment. *Academy of Management Journal*, 42, 58-74.
- KOÇEL, T. (2007).** İşletme Yöneticiliği. İstanbul: Arıkan Yayınları.
- KOYUNCUGIL, A. S. (2007).** “Veri Madenciliği ve Sermaye Piyasalarına Uygulaması”, Sermaye Piyasası Kurulu Araştırma Raporu, Araştırma Dairesi, ASK/1.
- KURT, I., TURE, M., & KURUM, T. (2008).** Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Systems with Applications*, 34, 366-374.
- LAI, R. K., FAN, C., HUANG, W. H., & CHANG, P. (2009).** Evolving and clustering fuzzy decision tree for financial time series data forecasting. *Expert Systems with Applications*, 36, 3761-3773.
- LEE, T. S., CHIUB, C., CHOUC, Y. C., & LUD, C. J. (2006).** Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis*, 50, 1113-1130.
- LEMERCIER, B., LACOSTE, M., LOUM, M., & WALTER, C. (2012).** Extrapolation at regional scale of local soil knowledge using boosted classification trees: A two-step approach. *Geoderma*, 171-172, 1-98.
- LEONG, T. Y. (1998).** Multiple perspective dynamic decision making. *Artificial Intelligence*, 105, 209-261.
- LI, Y. (2006).** Predicting materials properties and behavior using classification and regression trees. *Materials Science and Engineering*, 433, 261-268.
- LINDEN, R. C., WAYNE, S. J., & SPARROWE, R. T. (2000).** An examination of the mediating role of psychological empowerment on the relations between the job, interpersonal relationships, and work outcomes. *Journal of Applied Psychology*, 85, 407-416.
- MAHJOOBIA, J., & ETEMAD, S. (2008).** A. An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Applied Ocean Research*, 30, 172-177.

- MALAKOOTI, B. (2010).** Decision making process: Typology, intelligence, and optimization. *J Intell Manuf*, DOI 10.1007/s10845-010-0424-1.
- MALAKOOTI, B. (2011).** Systematic decision process for intelligent decision making, *J Intell Manuf*, 22, 627-642, DOI 10.1007/s10845-009-0327-1.
- MARTENS, R., VEALEY, R. S., & BURTON, D. (1990).** Competitive anxiety in sport. Champaign, IL: Human Kinetics.
- MARTINSONS, M. G., & DAVISON, R. M. (2007).** Strategic decision making and support systems: Comparing American, Japanese and Chinese management. *Decision Support Systems*, 43, 284-300.
- MOISEN, G. G. (2008).** Classification and regression trees. In Sven Erik Jørgensen and Brian D. Fath (Editor-in-chief), *Ecological informatics, Encyclopedia of ecology*, 1, (p. 582-588).
- NARCISS, S., KOERNDLE, H., & DRESEL, M. (2011).** Self-evaluation accuracy and satisfaction with performance: Are there affective costs or benefits of positive self-evaluation bias? *International Journal of Educational Research*. DOI:10.1016/j.ijer.2011.08.004.
- NEUMANNA, A., JOSIANE HOLSTEINA, J., LE GALLB, J. R., & LEPAGE, E. (2004).** Measuring performance in health care: Case-mix adjustment by boosted decision trees. *Artificial Intelligence in Medicine*, 32, 97-113.
- NIE, G., ROWE, W., ZHANG, L., TIAN, Y., & SHI, Y. (2011).** Credit card churn forecasting by logistic regression and decision tree. *Expert Systems with Applications*, 38, 15273-15285.
- NIEHOFF, B. P., MOORMAN, R. H., BLAKELY, G., & FULLER, J. (2001).** The influence of empowerment and job enrichment on employee loyalty in a downsizing environment, *Group & Organization Management*, 26, 93-113.
- NIELSEN, B. B., & NIELSEN, S. (2011).** The role of top management team international orientation in international strategic decision-making: The choice of foreign entry mode, *Journal of World Business*, 46, 185-193.
- OSEI-BRYSON, K.-M. (2007).** Post-pruning in decision tree induction using multiple performance measures, *Computers & Operations Research*, 34, 3331-3345.
- ÖZDEVECİOĞLU, M. (2003).** Algılanan Örgütsel Destek İle Örgütsel Bağlılık Arasındaki İlişkilerin Belirlenmesine Yönelik Bir Araştırma. *D.E.Ü.İ.İ.B.F Dergisi*, 18(2), 113-130.
- ÖZKAN, Y. (2008).** Veri Madenciliği Yöntemleri. İstanbul: Papatya Yayıncılık.
- PECCEI, R., & ROSENTHAL, P. (2001).** *DELIVERING* customer-oriented behaviour through empowerment: An empirical test of HRM assumptions. *Journal of Management Studies*, 38, 831-857.
- PODSAKOF, P. M., & MACKENZIE, S. B. (1997).** Impact of organizational citizenship behavior on organizational performance: A

- review and suggestions for future research. *Human Performance*, 10(2), 133-151.
- RAZI, M. A., & ATHAPPILLY, K. (2009).** Comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Systems with Applications*, 29, 65-74.
- ROWELL, T. A. (2009).** Management planning guidance for protected sites in the UK: A comparison of decision-making processes in nine guides. *Journal for Nature Conservation*, 17, 168-180.
- SEOL, H., CHOI, J., PARK, G., & PARK, Y. (2007).** A framework for benchmarking service process using data envelopment analysis and decision tree. *Expert Systems with Applications*, 32, 432-440.
- SHIN, Y., KIM, T., CHO, H., & KANG, K. (2012).** A formwork method selection model based on boosted decision trees in tall building construction. *Automation in Construction*, 23, 47-54.
- SORENSEN, E. H., MILLER, K. L., & OOI, C. K. (2000).** The decision tree approach to stock selection. *Journal of Portfolio Management*, 42-45.
- SPREITZER, G. M., KIZILOS, M. A., & NASSON, S. W. (1997).** A dimensional analysis of the relationship between psychological empowerment and effectiveness, satisfaction, and strain. *Journal of Management*, 23(5), 679-704.
- TANG, T. W., & TANG, Y. Y. (2011).** Promoting service-oriented organizational citizenship behaviors in hotels: The role of high-performance human resource practices and organizational social climates. *International Journal of Hospitality Management*, doi:10.1016/j.ijhm.2011.10.007.
- TOHIDI, H., & JABBARI, M. M. (2012).** Decision role in management to increase effectiveness of an organization. *Procedia - Social and Behavioral Sciences*, 31, 825-828.
- TUTEN, T. L., & NEIDERMEYER, P. E. (2004).** Performance, satisfaction and turnover in call centers: The effects of stress and optimism. *Journal of Business Research*, 57, 26-34.
- ÜNGÜREN, E., & DOĞAN, H. (2010).** Beş Yıldızlı Konaklama İşletmelerinde Çalışanların İş Tatmin Düzeylerinin CHAİD Analiz Yöntemiyle Değerlendirilmesi, *Cumhuriyet Üniversitesi İİBF Dergisi*, 11(2), 39-52.
- VEITCH, J. A., CHARLES, K. E., KELLY, M. J., FARLEY, K. M. J., & NEWSHAM, G. R. (2007).** A model of satisfaction with open-plan office conditions: COPE field findings. *Journal of Environmental Psychology*, 27, 177-189.
- WIDENER, S. K. (2006).** Human capital, pay structure, and the use of performance measures in bonus compensation. *Management Accounting Research*, 17, 198-221.
- WOODMAN, T., AKEHURST, S., HARDY, L., & BEATTIE, S. (2010).** Self-confidence and performance: A little self-doubt helps. *Psychology of Sport and Exercise*, 11, 467-470.

- WU, Z., & PAGELL, M. (2011).** Balancing priorities: Decision-making in sustainable supply chain management. *Journal of Operations Management*, 29, 577-590.
- YEH, W. Y., CHENG, Y., & CHEN, C. J. (2009).** Social patterns of pay systems and their associations with psychosocial job characteristics and burnout among paid employees in Taiwan. *Social Science and Medicine*, 68(8), 1407-1415.
- YERGE, N. E., & KARA, G. (2009).** Kriz Döneminde Kaliteden Kaçış, Eskişehir Osmangazi Üniversitesi, (Ödüllü Proje Çalışması).
- YOON M. H., & SUH J. (2003).** Organizational citizenship behaviours and service quality as external effectiveness of contact employees. *Journal of Business Research*, 56, 597-611.
- ZHANG, Z., WANG, D., & JIA, M. (2008).** Do high-performance human resource practices help corporate entrepreneurship? The mediating role of organizational citizenship behavior. *Journal of High Technology Management Research*, 19, 128-138.
- ZHAO, X., QU, H., & GHISELLI, R. (2011).** Examining the relationship of work-family conflict to job and life satisfaction: A case of hotel sales managers. *International Journal of Hospitality Management*, 30, 46-54.

ÇALIŞANLARIN PERFORMANSINA ETKİ EDEN DEĞİŞKENLERİN KARAR AĞAÇLARI YOLUYLA ANALİZİ VE FARKLI ALGORİTMALARIN KARŞILAŞTIRILMASI

Özet: Günümüz rekabet koşulları göz önüne alındığında farklı sektörlerde faaliyet gösteren örgütlerde çok sayıda değişkenle belirginleşen, ortaklaşan ya da ayrışan yönetsel problemler meydana gelmektedir. Bu problemlerin çözümünde örgütün bütününe yönelik değişiklik yapma, düzenleme, iyileştirme ve örgütsel geliştirme çalışmalarında çok sayıda değişkenin önem derecesine göre sınıflandırılmasının çeşitli problemlerin çözümünü kolaylaştırarak yönetsel etkinliği artırabilmesi mümkündür. Bununla birlikte Yönetim ve Organizasyon alanındaki araştırmalar incelendiğinde veri toplama yöntemi olarak genellikle ankete başvurulduğu ve belirli hipotezler çerçevesinde ilişkisel ve/veya etkiel tespitlerin yapıldığı görülmektedir. Katılanların görüş, tutum ve değerlendirmelerini irdeleme fırsatı veren anketlerin muhtemel bir sübjektif temele sahip olması mümkündür. Bu yüzden alışılmış analiz yöntemlerinin dışına çıkılması yönetim alanındaki araştırmacılara belirgin kolaylıklar sağlayabilir. Bu çalışmada veri madenciliğine ilişkin Regresyon Ağaçları (Classification Tree) ve Hızlandırılmış Ağaçlar (Boosted Decision Tree) yöntemleri kullanılarak, çalışanların algılanan performansına etki eden bazı örgütsel ve demografik değişkenlerin önem düzeyine göre sıralanması ve ağaç tabanlı algoritmaların örgütsel araştırmalarda kullanılabilirliğinin ortaya konulması amaçlanmaktadır. Karar Ağaçları ve ağaç tabanlı çeşitli algoritmalar veri madenciliğinin popüler ve etkili yöntemleri arasındadır. Literatüre bakıldığında daha çok tıp, mühendislik ve endüstriyel araştırmalarda kullanılan karar ağaçlarına yönelik algoritmaların sosyal bilimler ve özellikle yönetim ve organizasyon disiplini içerisinde kullanımının oldukça yeni olduğu söylenebilir. Ağaç tabanlı algoritmalar, istatistiksel yöntemlere göre örgütsel değişkenler arasındaki ilişkinin yönünü ve düzeyini görsel bir ortamda sunabilme özelliğine sahiptir. Karar ağaçları tekniğinin yönetim ve organizasyon kapsamında yaygınlaşması örgütsel araştırmalarda daha somut ve basit yorumların yapılarak hızlı ve doğru karar alabilme avantajını pekiştireceği söylenebilir. Böylelikle yöneticilerin işe alım sürecinden itibaren hangi adayların istenilen niteliklere uygun olduğuna yönelik rasyonel kriterler oluşturulabilmesi ve farklı çalışan gruplarına uygun örgütsel strateji, politika ve taktikler geliştirebilmesi mümkün hale gelmektedir. Araştırmanın örneklemini Kırşehir ilinde faaliyet gösteren ve çalışan sayısı 50'nin üzerinde olan 10 özel sektör işletmesinin çalışanlarından oluşmaktadır. Anket formuna ilişkin geçerlilik ve güvenilirlik testleri yapılmış olup, analizde DTREG paket programı kullanılmıştır. Araştırmadan elde edilen bulgulara göre örgütsel vatandaşlık, iş tatmini, yaşam tatmini ve çalışılan pozisyon çalışanların performansına etki eden en önemli değişkenler olarak belirlenirken, Hızlandırılmış Ağaçlar algoritmasının daha tutarlı ve güvenilir sonuçlar meydana getirdiği tespit edilmiştir. Karar Ağaçları tekniği örgütsel değişkenler arası ilişkiler ve değişkenlerin önem düzeylerinin belirlenmesinde alternatif bir yol olarak kullanılabilir. Ancak (örgütsel araştırmalarda karar ağaçları tekniğinin) güven düzeyinin yükseltilebilmesi için Regresyon Ağaçları yöntemi ile bu yöntemin geçerliliğini artırabilen Hızlandırılmış Ağaçlar algoritmasının birlikte kullanılması tavsiye edilebilir. Zira Regresyon Ağaçları görsel bir ağaç sunabilirken, Hızlandırılmış Ağaçlar daha yüksek güven ve tutarlı sonuçlarla birlikte 100'lerce işlev ve dallanma sebebiyle görsel sonuçlar sunamamaktadır. Nihai olarak Hızlandırılmış Ağaçlar algoritması, Regresyon Ağaçları gibi bazı algoritmalarda öngörüsül doğruluğu artırmak için rahatlıkla kullanılabilir.

Anhtar Kelimeler: Yönetim ve Organizasyon, Performans, Veri Madenciliği, Karar Ağaçları