Decision Tree Induction of Emotional Violence against Women

Esra Kahya Özyirmidokuz[1,a,*] and Yasemin Kaya[2,b]

[1] Kayseri Vocational College, Erciyes University, Kayseri, Turkey

[2]Social Science Vocational College, Erciyes University, Kayseri, Turkey

[a]esrakahya@erciyes.edu.tr, [b]ykaya@erciyes.edu.tr

*Corresponding author

Keywords: Information management, Data mining, Decision trees, Emotional violence.

Abstract. Knowledge Discovery and Data Mining (KDD) is an interdisciplinary area focusing upon methodologies for extracting useful and hidden knowledge from data. KDD is the process of analyzing data from different perspectives and summarizing them into new information. It uses various types of algorithms, including statistical and machine learning. KDD starts with understanding data. Then data preprocessing techniques are used because quality decisions must be based on quality data. The data mining modeling phase is used to extract more meaningful knowledge from a given data set. Decision trees (DTs) are one of the most powerful and popular approaches in data mining. Their effective structure helps us to make the best decisions on the basis of existing information. DTs also predict future and possible decisions.

Analyzing data on violence against women is a very important subject. There are still a lot of women who are subjected to violence, even in this, the 21st century. Violence against women is not only hysical; emotional violence also negatively affects health. Emotional violence inflected by a partner includes the following: being insulted or made to feel bad about oneself; being humiliated in front of others; being intimidated or frightened on purpose; being threatened directly, or indirectly through a threat to someone the respondent cares about. Although violence against women, especially by men, is an important issue, presently there are too few original studies on this topic. Women who experience emotional violence are at increased risk of injury and death, as well as a range of physical, emotional, and social problems. As women play an important role in bringing up healthy individuals to be part of a healty community, violence against women is also a problem for society. This research aims to extract meaningful knowledge from the emotional violence against women data set. The data cover 12795 women in the 15-59 age group of in Turkey. The data were obtained from the Turkish Statistical Institute (TURKSTAT)'s Web site. As a result of the data obtained from the analysis suggestions can be submitted for women exposed to emotional violence.

The paper explores the use of different preprocessing and data mining techniques (anomaly detection analysis, data reduction, data cleaning, DT etc.). In order to support the prevention of violence against women, a DT model is generated. Modeling is performed with C4.5, which is one of the popular DT modeling algorithms and an extension of the earlier well known ID3 algorithm to find useful patterns.

The cross-validation method involves partitioning the examples randomly into n folds. Ten fold normal cross-validation accuracy evaluation is used.

Consequently, a DT is generated. The resulting DT model can be used for future decisions. The relationships of the input variables with target variables are found. The results show that the proposed method has a substantially good performance.

ISBN: 978-605-64453-0-9

Introduction

The violence that we can face in every field of life is defined by the World Health Organization as "the intentional use of physical force or power, threatened or actual, against oneself, another person, or against a group or community, that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment, or deprivation."

There are many studies on violence against women. Physical force or power and identifying the result or effect of it is relatively easy. However, emotional violence and the result of it cannot be identified easily. Although it is difficult to make certain judgements about these issues in the field of social science, there are no research methods to support research. DT is used in this study to aid the analysis of the emotional violence against women. The model is evaluated with data about domestic violence.

Data mining (DM) uses many disciplines including statistics, machine learning, databases, data visualization, information management ect. It extracts important, previously unknown relations from data. DM versus classical techniques is used to improve performance and better understand data. DTs are popular DM techniques. One of the greatest advantages of DTs is the fact that they are useful for collapsing a set of categorical values. Therefore DTs are suitable for the analysis of social issues. However, there is no research about emotional violence against women via data mining or DTs.

DM techniques are being increasingly used in many modern organizations to retrieve valuable knowledge structures from organizational databases, including data warehouses. An important knowledge structure that can result from data mining activities is the DT, which is used for the classification of future events. The induction of the DT is done using a supervised knowledge discovery process in which prior knowledge regarding classes in the database is used to guide the discovery. The generation of a DT is a relatively easy task but in order to select the most appropriate DT it is necessary for the DM project team to generate and analyze a significant number of DTs based on multiple performance measures [1].

DTs are used in business, information, marketing, finance, philosophy, decision analysis etc. DM studies have been carried out on social issues in recent years [2,3,4]. There are many different applications of DM in the social sciences, such as in process planning and management[5], tourism management[6], finance [7], marketing [8,9], quality control and manufacturing [10, 11] procurement and knowledge management [12,13,14].

DTs are graphical representations of alternative choices, which enable the decision maker to identify the most suitable option. DTs have been in use for over 50 years [15,16, 17,18,19,20,21,22]. Kim and Koehler [23] summarized various theoretical results of pruning a DT and showed their usefulness. Empirical results have shown that pruning a DT sometimes improves its accuracy.

Several methods used hybrid DTs at the begining of the 2000s. Setiono and Leow [24] described a pruning-based method for mapping DTs to neural networks, which could compress the network by removing unimportant and redundant units and connections. Zhou and Chen [25] proposed a hybrid DT which simulated human reasoning by using symbolic learning to do qualitative analysis and neural learning to do the subsequent quantitative analysis. Osei-Bryson [1] proposed a multi-criteria decision analysis based process that would empower DM project teams to conduct thorough experimentation and analysis without being overwhelmed by the task of analyzing a significant number of DTs would make a positive contribution to the DM process.

Larose [26] used the cluster memberships in DM. A CART DT model was run to classify customers as either churners or nonchurners. Lye et al. [27] used Back-propagation Neural Network (BPNN), Classification and Regression Tree (CART), and Generalized Regression Neural Network (GRNN) in predicting students’ mathematics achievement to predict the students’ mid-semester evaluation results, whereas the latter part employed additional data to predict students’ final examination results. The predictive models’ accuracy was evaluated by using 10-fold cross-validation to identify the best model. The findings revealed that BPNN outperformed other models with an

ISBN: 978-605-64453-0-9

accuracy of 66.67% and 71.11% in predicting the mid-semester evaluation result and the final examination result respectively.

Altıncay [28], proposed the use of model ensemble-based nodes where a multitude of models were considered for making decisions at each node. The ensemble members were generated by perturbing the model parameters and input attributes. Experiments conducted on several datasets and three model types indicated that the proposed approach achieved better classification accuracies compared to individual nodes, even in cases when only one model class was used in generating ensemble members. Wei et al. [29], presented a new approach for inducing DTs based on the Variable Precision Rough Set Model. The presented approach was aimed at handling uncertain information during the process of inducing DTs and generalized the rough set based approach to DT construction by allowing some degree of misclassification when classifying objects. Lee et al. [30], focused on the acceptability of e-commerce, which is a prerequisite for offline service providers to succeed in the online marketplace, as a business medium for services. They developed a DT model for predicting whether or not e-commerce would be successful for a specific type of service.

Aitkenhead [31] presented an evolutionary method which allowed DT flexibility through the use of co-evolving competition between the DT and the training data set. This method was tested using two different datasets and gave results comparable with or superior to other classification methods. Abrahams et al. [32] demonstrated the use of DT induction for the creation of a marketing strategy for a new pet insurance company, PetPlan USA. They employed both a traditional C4.5 DT approach, and a novel locally profit-optimal decision algorithm, called SBP, to discover the characteristics of profitable demographics for PetPlan to market to. Zengin et al. [33] presented a sample study analyzing data gathered from an educational study using DM techniques appropriate for processing these data. T-test, analysis of variance and DTs were used, as well as dependency networks and clustering. Diskaya et al. [34] analyzed a variety of economic signs by using artificial neural networks and DTs algorithms. Rojas et al. [35] presented a proposal of representation and scheme of exploratory visualization for DTs in the KDD (Knowledge Discovery in Database) process, specifically in the DM stage. Azad et al. [36] presented three approaches for the construction of decision rules for decision tables with many-valued decisions.

This research aims to extract meaningful knowledge from data on emotional violence against women. The subject of previous works about women has generally been on mobbing or physical violence in the family. One of the differences of this research is that it is not about physical domestic violence. It uses emotional violence data to extract a DT. The findings from this study make several contributions to the current literature.

The rest of the paper is organized as follows. The next section presents a brief description of the DM and DTs. In Section **3**, the model is presented including the preprocessing techniques. In Section **4** we draw our conclusions.

1. Data mining and decision trees

Data mining (DM) is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques. There are other definitions [26]:

- DM is the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner.
- DM is an interdisciplinary field bringing together techniques from machine learning, pattern recognition, statistics, databases, and visualization to address the issue of information extraction from large databases.

ISBN: 978-605-64453-0-9

A lot of data are currently being collected. However, what is being learned from all these data? What knowledge do we gain from all this information? Probably depending on the supermarket, not much. The ongoing remarkable growth in the field of DM and knowledge discovery has been fueled by a fortunate confluence of a variety of factors [26]:

- The explosive growth in data collection, as exemplified by the supermarket scanners above
- The storing of data in data warehouses, so that the entire enterprise has access to a reliable current database
- The availability of increased access to data from Web navigation and intranets
- The competitive pressure to increase market share in a globalized economy
- The development of off-the-shelf commercial DM software suites
- The tremendous growth in computing power and storage capacity

Considerable attention has been devoted by the machine learning research community to the task of acquiring "classification knowledge" for which, among a predeclared set of available classes, the objective is to choose the most appropriate class for a given case. The goal in such research is to develop methods that induce the desired classification knowledge from a given set of preclassified examples. Significant progress has been made in the last decade toward this goal, and various methods for automatically inducing *classifiers* from data are now available. In particular, constructing classifiers in the form of DTs has been quite popular, and a number of successful real-world applications that employ DT construction methods have been reported [37].

DTs are a form of multiple variable (or multiple effect) analyses. All forms of multiple variable analyses allow us to predict, explain, describe, or classify an outcome (or target). An example of a multiple variable analysis is a probability of sale or the likelihood to respond to a marketing campaign as a result of the combined effects of multiple input variables, factors, or dimensions. This multiple variable analysis capability of DTs enables us to go beyond simple one-cause, one-effect relationships and to discover and describe things in the context of multiple influences. Multiple variable analysis is particularly important in current problem-solving because almost all critical outcomes that determine success are based on multiple factors. Further, it is becoming increasingly clear that while it is easy to set up one-cause, one-effect relationships in the form of tables or graphs, this approach can lead to costly and misleading outcomes [38].

DM algorithms identify valid, novel, potentially useful, and ultimately understandable patterns from data that can be used for making high confidence classifications. A typical DM algorithm generates rules that describe relationships between the input features and an outcome. Discovering hidden patterns in the data may represent valuable knowledge that might lead to discoveries. There are various DM algorithms ranging from decision trees (DTs) to clustering [39].

DTs are one of the popular methods to have been used for knowledge discovery in databases. Tree models can be defined as a recursive procedure, through which a set of statistical units is progressively divided into groups, according to a division of an explanatory variable to split and the choice of a splitting rule for such a variable, which establishes how to partition the observations. The main result of a tree model is a final partition of the observations. To achieve this it is necessary to specify stopping criteria for the division process [40]. In DM, a DT is a predictive model which can be used to represent both classifiers and regression models.

Although DTs are powerful and popular tools for classification and prediction, they need an expert to interpret and understand them. To express DTs, decision rules are used. DTs are graphic tools that represent rules. DT induction is closely related to rule induction. Each path from the root of a DT to one of its leaves can be transformed into a rule simply by conjoining the tests along the path to form the antecedent part, and taking the leaf's class prediction as the class value [41].

As DTs evolved, they turned out to have many useful features, both in the traditional fields of science and engineering and in a range of applied areas, including business intelligence and DM. These features include the following [38] :

- DTs produce results that communicate very well in symbolic and visual terms. DTs are easy to produce, easy to understand, and easy to use. One useful feature is their ability to incorporate multiple predictors in a simple, step-by-step fashion. The ability to incrementally build highly complex rule sets (which are built on simple, single association rules) is both simple and powerful.
- DTs readily incorporate various levels of measurement, including qualitative (e.g., good-bad) and quantitative measurements. Qualitative measurements include ordinal (e.g., high, medium, low categories) and interval (e.g., income, weight ranges) levels of measurement.
- DTs readily adapt to various twists and turns in data –unbalanced effects, nested effects, offsetting effects, interactions and nonlinearities- that frequently defeat other one-way and multi-way statistical and numeric approaches.
- DTs are nonparametric and highly robust (for example, they readily accommodate the incorporation of missing values) and produce similar effects regardless of the level of measurement of the fields that are used to construct DT branches (for example, a DT of income distribution will reveal similar results regardless of whether income is measured in 000s, in 10s of thousands, or even as a discrete range of values from 1 to 5).

A typical DT is shown in Fig.1. It represents the concept buys_computer, that is, it predicts whether or not a customer at AllElectronics is likely to purchase a computer. Each internal (nonleaf) node represents a test on an attribute. Each leaf node represents a class (either buys_computer=yes or buys_computer=no). Internal nodes are denoted by rectangles, and leaf nodes are denoted by ovals. In order to classify an unknown sample, the attribute values of the sample are tested against the DT. A path is traced from the root to a leaf node that holds the class prediction for that sample. DTs can easily be converted to classification rules [42].

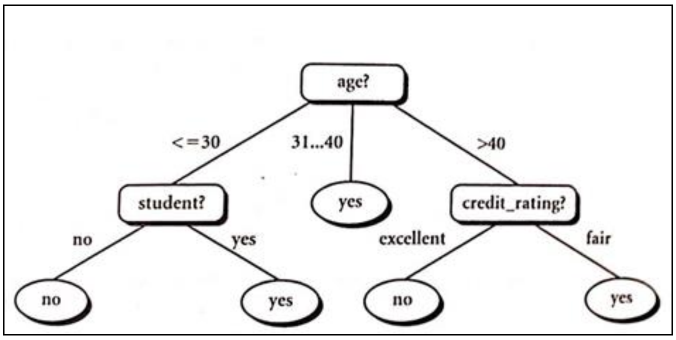

Fig.1. A DT for the concept buys_computer

For knowledge-based systems, DTs have the advantage of being comprehensible to human experts and of being directly convertible into production rules. Moreover, when used to handle a given case, a DT not only provides the solution for that case, but also states the reasons behind its choice. These features are very important in typical application domains in which human experts seek tools to aid in conducting their job while remaining "in the driver's seat." Another advantage of using DTs is the ease and efficiency of their construction compared to that of other classifiers such as neural networks [37].

ISBN: 978-605-64453-0-9

2. Application

This study uses TUIK's violence against women research data which were collected from 22,822 women in households in 2008. Researchers interviewed 12,795 women. Women's questionnaire response rate was 86.1%. The data cover 12,795 women in the 15-19 age group in Turkey. The data were collected from the Turkish Statistical Institute (TURKSTAT)'s Web site. As a result of the data obtained from the analysis suggestions can be submitted for women exposed to emotional violence [43]. Fig.2. shows the knowledge discovery process of the application.

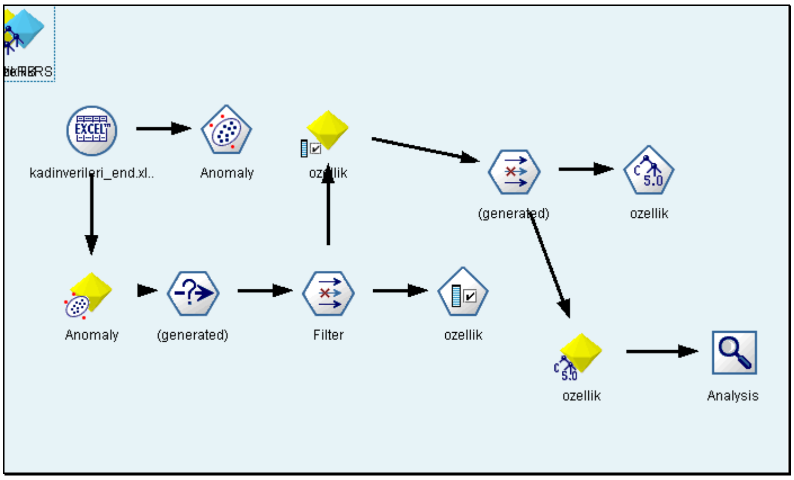

Fig.2. A screenshot of SPSS Clementine programme

2.1. Data preprocessing

Data preparation consists of a diverse set of processes to clean and transform the data. We use an anomaly detection model to identify outliers, or unusual cases, in the data. Unlike other modeling methods that store rules about unusual cases, anomaly detection models store information on what normal behavior looks like. This makes it possible to identify outliers even if they do not conform to any known pattern, and it can be particularly useful in applications. While traditional methods of identifying outliers generally look at one or two variables at a time, anomaly detection can examine large numbers of fields to identify clusters or peer groups into which similar records fall. Each record can then be compared to others in its peer group to identify possible anomalies. The further away a case is from the "normal" center, the more likely it is to be unusual. Clementine calculates an anomaly index for each record which is the ratio of the group deviation index to its average over the cluster that the case belongs to [13]. Two groups are detected. Consequently, those records which are greater than the average anomaly index level (0.97626) are eliminated.

We use feature selection analysis to eliminate the redundant attributes. We eliminate 1 attribute after the feature selection model.

2.2. Modeling with decision trees

In order to support the prevention of violence against women, a DT model is generated. Modeling is performed with C4.5, which is one of the popular DT modeling algorithms and an extension of the earlier well known algorithm ID3, to find useful patterns. Nine attributes are used in modeling phase. A C4.5 tree with four tree depths is generated. The C4.5 DT model can be seen from Fig. 3.

ISBN: 978-605-64453-0-9

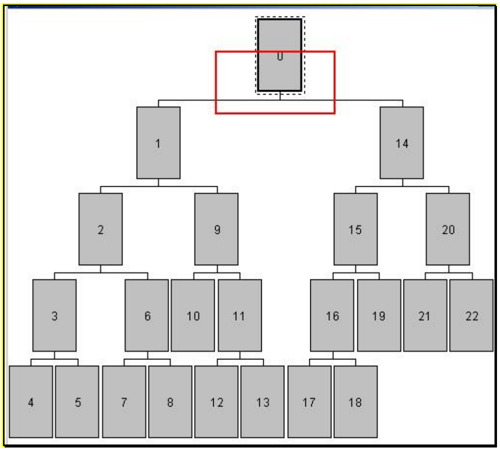

Fig.3. The DT model

DTs are tree shaped structures that represent sets of decisions. They are useful because they provide strategic answers for uncertain problems with a graphical representation of how different factors may influence the whole model. We generate a C4.5 DT from violence against women data. The DT approach can generate rules for the classification of a data set. The generated rules can be seen in Fig.4.

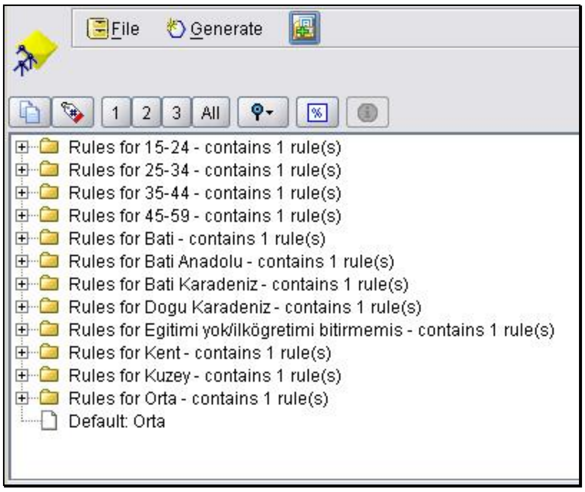

Fig.4. The rule induction

The following examples illustrate some of the extracted rules:

- If C1=true and C2=false and C3=true and C4>36 then "bati"
- If C1=false and C2=true and C3<24 then "Egitimi yok/ilkokulu bitirmemis"

2.3. Accuracy

The accuracy of a classifier refers to the ability of a given classifier to correctly predict the class label of new or previously unseen data (i.e., tuples without class label information). Similarly, the accuracy of a predictor refers to how well a given predictor can guess the value of the predicted attribute for

ISBN: 978-605-64453-0-9

new or previously unseen data. Accuracy can be estimated using one or more test sets that are independent of the training set. Cross-validation and bootstrapping are some of estimation techniques used [39].

The cross-validation method involves partitioning the examples randomly into *n* folds. (Ten is a fairly popular choice for *n*, but much depends on the number of examples available.) We use one partition as a testing set and use the remaining partitions to form a training set. As before, we apply an algorithm to the training set and evaluate the resulting model on the testing set, calculating the percentage correctly. We repeat this process by using each of the partitions as the testing set and using the remaining partitions to form a training set. The overall accuracy is the accuracy averaged over the number of runs, which is equivalent to the number of partitions. *Stratified* cross-validation involves creating partitions so that the number of examples of each class is proportional to the number in the original set of examples [44]. The accuracy ratio of the model is 66.67%. Ten-fold normal cross-validation accuracy evaluation is used.

3. Conclusions

Violence against women is a condition that can have a negative impact on society. Domestic violence and what women's rights are in this regard have produced no information that is valuable in this respect. An increase in public awareness is necessary in order to try to solve this problem before it occurs rather than after it has occurred. The increase in the level of awareness on this issue is directly related to the information produced. Violence against women during the process of socialization in disadvantaged groups could be a conducted for a change of perspective. Domestic violence is related sociology, psychology, behavioral science and law. It requires an interdisciplinary perspective. Hence, a DT is extracted in this study to produce valuable information.

4. References

[1] K.-M. Osei-Bryson, Evaluation of DTs: a multi-criteria approach, *Computers & Operations Research* 31 (2004) 1933–1945.

[2] G. Ritschard, A. Gabadinho, N. S. Muller and M. Studer, Mining event histories: a social science perspective, Int. J. Data Mining, *Modelling and Management*, Vol. 1, No. 1, pp. 1-23(preprint) 1, 2009.

[3] A. Buldu and K. Üçgün, Data mining application on students' data, *Procedia Social and Behavioral Sciences* 2, pp. 5251–5259, 2010.

[4] M. Chalaris, I. Chalaris, C. Skourlas and A. Tsolakidis, The 2nd International Conference on Integrated Information, Extraction of rules based on students' questionnaires, Procedia - Social and Behavioral Sciences 73, pp. 510 – 517, 2013.

[5] H. Seol, J. Choi and G. Park, Yongtae Park A framework for benchmarking service process using data envelopment analysis and DT, *Expert Systems with Applications* 32, pp. 432–440, 2007.

[6] S.S. Kim, D. J. Timothy and J. Hwang, Understanding Japanese tourists' shopping preferences using the DT Analysis method, *Tourism Management* 32, pp. 544-554. 2011.

[7] P. Hildebrandt and T. Knoke, Investment decisions under uncertainty—A methodological review on forest science studies, *Forest Policy and Economics* 13, pp. 1–15, 2011.

[8] K. Yada, E. Ip and N. Katoh, Is this brand ephemeral? A multivariate tree-based decision analysis of new product sustainability, *Decision Support Systems* 44, pp.223–234, 2007.

[9] D.L. Olson and B. K. Chae, Direct marketing decision support through predictive customer response modeling, *Decision Support Systems,* 54, pp.443–451, 2012.

[10] C. Wua and D. Barnes, A literature review of decision-making models and approaches for partner selection in agile supply chains, *Journal of Purchasing & Supply Management* 17, pp. 256–274, 2011.

[11] C. Çiflikli and E. Kahya-Özyirmidokuz, Enhancing product quality of a process, *Industrial Management & Data Systems*, vol.112, no. 8, pp.1181-1200, 2012.

[12] S. Lee, S. Lee and Y. Park , A prediction model for success of services in e-commerce using DT: E-customer's attitude towards online service, *Expert Systems with Applications* 33, pp. 572–581, 2007.

[13] C. Çiflikli and E. Kahya-Özyirmidokuz, Implementing a DM solution for enhancing carpet manufacturing productivity, *Knowledge Based Systems*, vol.23, pp.783-788, 2010.

[14] G. Nie, W. Rowe, L. Zhang, Y. Tian and Y. Shi, Credit card churn forecasting by logistic regression and DT, *Expert Systems with Applications* 38, pp. 15273–15285, 2011.
[15] G.V. Kass, An Exploratory Technique for Investigating Large Quantities of Categorical Data, *Applied Statistics,* Vol. 29 No.2, pp. 119-127. 1980.
[16] L. Breiman, J.H. Friedman, R.A. Olshen and C.J. Stone, *Classification and Regression Trees*, Chapman and Hall/CRC. 1984.
[17] I. Kononenko, I. Bratko, and E. Roskar, Experiments in automatic learning of medical diagnostic rules, *Technical report*, Josef Stefan Institute, Ljubljana, Yugoslavia, 1984.
[18] J.R. Quinlan, Induction of DTs, Machine Learning, Vol.1, pp.81-106, 1986.
[19] J.R. Quinlan, *C4.5: Programs for machine learning*. Morgan Kaufmann, 1993.
[20] D. Michie, Learning concepts from data, *Expert Systems with Applications* 15, pp. 193–204, 1998.
[21] D. McSherry, Strategic induction of DTs, *Knowledge-Based Systems* 12, pp. 269–275, 1999.
[22] S. Piramuthu, Input data for DTs, *Expert Systems with Applications* 34, pp.1220–1226, 2008.
[23] H. Kim, G.J. Koehler, Theory and Practice of DT Induction, Int. J. MgmtSci., Vol. 23, No. 6, pp. 637-652, 1995.
[24] R. Setiono and W.K. Leow, On mapping DTs and neural networks, *Knowledge-Based Systems* 12 (1999) 95–99.
[25] Z-H, Zhou, Z.-Q. Chen, Hybrid DT, *Knowledge Based Systems*, vol.15, pp.515-528, 2002.
[26] D.T. Larose, Discovering knowledge in data: an introduction to DM, Wiley, 2005.
[27] C.-T. Lye, L.-N. Ngb, M. D. Hassanc, W.-W.Gohd, Predicting pre-university student's mathematics achievement, *Procedia - Social and Behavioral Sciences*, vol. 8, pp. 299-306, 2010.
[28] H. Altıncay, DTs using model ensemble-based nodes, *Pattern Recognition* 40, pp. 3540 – 3551, 2007.
[29] J.M. Wei, S. Q. Wang, M.Y. Wang, J.P. You, D.Y. Liu, Rough set based approach for inducing DTs, *Knowledge-Based Systems* 20, pp. 695–702, 2007.
[30] S. Lee, S. Lee, Y. Park, A prediction model for success of services in e-commerce using DT: E-customer’s attitude towards online service, *Expert Systems with Applications*, vol. 33, pp.572–581, 2007.
[31] M.J. Aitkenhead, A co-evolving DT classification method, *Expert Systems with Applications* 34, pp.18–25, 2008.
[32] A.S. Abrahams, A.B. Becker, D. Sabido, R. D’Souza, G. Makriyiannis and M. Krasnodebski, Inducing a marketing strategy for a new pet insurance company using DTs, *Expert Systems with Applications* 36, pp.1914–1921. 2009.
[33] K. Zengin, N. Esgi, E. Erginer and M.E. Aksoy, A sample study on applying data mining research techniques in educational science: developing a more meaning of data, *Procedia Social and Behavioral Sciences,* 15, pp. 4028–4032, 2011.
[34] F. Diskaya, S. Emir and N. Orhan, Determining the Necessary Criteria for the EU Membership by Using the Machine Learning Method: 2005 - 2010 Period Analysis, Procedia Social and Behavioral Sciences 24 pp. 808–814, 7th International Strategic Management Conference, 2011.
[35] W.A.C. Rojas, C. J. M. Villegas, Graphical Representation and Exploratory Visualization for DTs in the KDD Process, The 2nd International Conference on Integrated Information, *Procedia - Social and Behavioral Sciences*, vol.73, pp.136 – 144, 2013.
[36] M. Azad, B. Zieloskoa, M. Moshkova and I. Chikalova, 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems - KES2013 Decision Rules, Trees and Tests for Tables with Many-Valued Decisions – Comparative Study, Procedia Computer Science 22 (2013) 87 – 94.
[37] H. Almuallım, S.Kaneda and Y. Akıba, Development and applications of DTs, *Expert Systems*, Edited by Cornelius T. Leondes, by Academic Press, Volume 1, pp.53- 77, 2002.
[38] B. de Ville, DTs for business intelligence and DM: Using SAS Enterprise Miner, SAS Institute Inc., USA 2006.
[39] A. Burns, A. Kusiak and T. Letsche, *Mining Transformed Data Sets,* In: M.G. Negoita et al. (Eds.), KES 2004, LNAI 3213, Springer-Verlag Berlin Heidelberg, pp. 148-154. 2004.
[40] P. Giudici, *Applied data mining.* Wiley, England, 2003.
[41] L. Rokach, O. Maimon, *Data mining with DTs.* World Scientific, USA, 2007.
[42] J. Han, M. Kamber**,** Data Mining : Concepts and Techniques, 2nd Ed., Elsevier, USA, 2006**.**
[43] Turkish Statistical Institute (TURKSTAT) Web Site, http://www.tuik.gov.tr.
[44] M.A. Maloof, Machine Learning and Data Mining for Computer Security, Methods and Applications, Springer-Verlag, USA. 2006.

ISBN: 978-605-64453-0-9