THE CLASSIFICATION OF THE PROBABILITY UNIT ABILITY LEVELS OF THE ELEVENTH GRADE TURKISH STUDENTS BY CLUSTER ANALYSIS

Assist. Prof. Dr. Ozcan OZYURT
Karadeniz Technical University,
Technology Faculty,
Software Engineering Department,
Trabzon, TURKEY

ABSTRACT

In this study, the probability unit ability levels of the eleventh grade Turkish students were classified through cluster analysis. The study was carried out in a high school located in Trabzon, Turkey during the fall semester of the 2011-2012 academic years. A total of 84 eleventh grade students participated. Students were taught about permutation, combination, binomial expansion, and probability, which were the sub-topics of probability unit, in an individualized mathematics learning environment called UZWEBMAT. After students completed the learning of each sub-topic, they were subjected to an exam about the relevant topic through UZWEBMAT-CAT. Students participated in 5 separate exams (i.e. one for each sub-topic and one end-of-unit test). Data were collected via system records made up of the ability levels of students concerning each subject. The ability levels obtained from each exam were analyzed through hierarchical clustering. According to the study results, the ability levels of students gathered in two main clusters in every test: medium ability level and advanced ability level.

Keywords: **Computerized Adaptive Testing, Individual Differences, Ability Level, Hierarchical Cluster Analysis.**

INTRODUCTION

Today's societies face big heaps of information. Meaningful and beneficial data should be extracted from such heaps of information. The extraction of meaningful data form big heaps of information is referred to as data mining. In the most general sense, data mining is known as the extraction of implicit patterns from big data sets (Klosgen & Zytkow, 2002; Romero & Ventura, 2007).

It is possible to observe data mining in many fields including education, health, banking, and e-commerce. The concept of Educational Data Mining (EDM) has emerged as a result of the extension of data mining applications over educational data. EDM is defined as the process of discovering meaningful patterns through educational data (Baker & Yasef, 2009; Lee, Chen, Chrysostomou, & Liu, 2009; Wang & Liao, 2011). The methods employed by EDM are as follows: statistics and visualization, clustering, classification, outlier, association, predication, and pattern matching (Baker & Yasef, 2009; Kotsiantis, Patriarcheas, Xenos, 2010; Levy & Wilensky, 2011).

EDM can be employed for evaluating the learning performances of students, ameliorating learning processes, guiding students' learning, giving feedbacks and adapting learning recommendations in accordance with the learning behaviors of students, evaluating learning materials and courseware, detecting abnormal learning behaviors and problems, and achieving a deeper understanding of educational phenomena (Baepler & Murdoch, 2010; Baker & Yasef, 2009; Chang, 2006; Chen, Hsieh, & Hsu, 2007; Lazcorreta, Botella, & Fernández-Caballero, 2008; Lee et al., 2009; Levy & Wilensky, 2011; Nandeshwar, Menzies, & Nelson, 2011; Romero & Ventura, 2010; Romero, Ventura, & Garcia, 2008).

Many studies have been carried out in the field of EDM in recent years. These studies have been conducted over the data acquired from traditional classroom environments and computer/web-aided learning environments (Chen & Liu, 2011; Lee, 2012; Mostow & Beck, 2006; Romero, Ventura, & Garcia, 2008; Tsantis & Castellani, 2001; Vandamme, Meskens, & Superby, 2007; Zafra, Romero, & Ventura, 2011). Most of these studies have been carried out through computer/web-aided learning environments. This is because all knowledge of students including actions and interactions can be recorded via databases and logs in computer/web-aided learning environments (Abdous & He, 2011; Romero, Ventura, & Garcia, 2008). The abundance of data acquired through computer/web-based learning environments has led to the variation of EDM applications carried out in these environments. For this reason, educational data mining is increasing its importance as a research area attempting to make use of the abundant data generated by various educational systems for improving teaching, learning and decision making (Baker & Yacef, 2009; Garcia, Romero, Ventura, & de Castro, 2011; He, 2013; Liao, Chu, & Hsiao, 2012).

In this study, exams were conducted on permutation, combination, binomial expansion, and probability, which were among the sub-topics of the probability unit, through the Computerized Adaptive Testing (UZWEBMAT-CAT) module integrated into UZWEBMAT (XXX, 2013). CAT systems adapt the difficulty levels of questions in accordance with the ability levels of individuals and yield highly precise measurement results (Kreitzberg, Stocrisg & Swansos, 1978; Weiss, 1985). UZWEBMAT-CAT calculates the knowledge levels of students in the range of -3 and +3. In this study, the ability levels obtained via exams were classified. Classification was performed based on hierarchical clustering method. In this way, the ability level intervals where the probability unit ability levels of students concentrated were determined.

The structure of this paper is organized as follows: Section 2 deals with the studies on EDM by use of the data acquired from web-based learning environments. Section 3 presents the details of research methodology. Section 4 describes the findings of the present study. Section 5 is about the results of the study.

RELATED WORKS

Researchers engaged in the field of EDM have carried out many studies on subjects such as individual learning; computer supported collaborative learning, and computerized adaptive testing (Baker & Yasef, 2009). Among these studies, the recently featured ones are as follows: Pal (2012) made an attempt to predict the engineering students who were likely to drop out in the first year, and used such classification algorithms as ID3, C4.5, CART and ADT decision tree over the data related to the students dropping out in their first years in previous periods.

According to the results of that study, the reasons of new-comers for dropping out are predicted in high accuracy by use of the data of previous students dropping out in their first years. Jovanovica, Vukicevica, Milovanovica, & Minovica (2012) grouped students in an e-learning environment based on their cognitive styles, and classified their performances. According to the research results, the fact that the students categorized based on their cognitive styles received materials suitable for themselves had a positive effect on their performances. Romero, Espejo, Zafra, Romero, & Ventura (2010) carried out some experimental students on the Moodle e-learning system. They demonstrated how the final exam grades of university students could be estimated through web mining applications over the Moodle. In addition, the researchers determined students with similar characteristics and students with low motivation by using classification algorithms. He (2013) employed data mining and text mining techniques in order to search the patterns of participation and interaction of students in a live video streaming environment through the examination of the data automatically acquired by the live video streaming environment.

In that study, 114 course data covering various subjects from computer sciences and 1144 student data were used. It was concluded that students from different departments had different interactions. Furthermore, a positive correlation was found between the interaction frequencies and achievements of students interacting with instructor. Falakmasir & Jafar (2010) utilized data mining in an attempt to rank students' activities that influenced their performances, which was measured based on their final grades. They concluded that the participation of students in virtual classrooms yielded the highest effect on their final grades. Romero et al. (2008) conducted data mining through student data on the Moodle e-learning system. In that study, researchers demonstrated how useful data mining applications could be for instructors. Lee et al. (2009) carried out a data mining process in order to determine the preferences of students in a web-based learning environment. Decision tree was used as an instrument of classification in the study which was conducted with 65 university students. According to the results of that study, cognitive style is an important factor that determines the preferences of students.

Moreover, the study revealed that decisions trees were quite beneficial for the classification of students according to their cognitive styles. Zafra et al. (2011) used the data of university students on the Moodle system. These data had been acquired from the quizzes, assignments, and forum activities of students.

In that study, the effect of these activities on student learning was studied through data mining applications. That study made an attempt to predict the performances of students. The main focus of this study was to explore whether data mining technology could be more effective in solving that problem using representation based on multiple instances rather than classical representation making use of single instances. Experimental results demonstrated how their representation based on multi instance learning was more effective and acquired more accurate models besides a more optimized representation, which eliminated the shortcomings of classical representation. Fausett & Elwasif (1994) predicted the grades of students from test scores via two types of neural networks: back propagation and counter propagation. According to the results of experimental studies, the highly rapid training of the counter propagation networks still makes them appealing alternative to back propagation for applications in which moderate accuracy is acceptable. Minaei-Bidgoli & Punch (2003) classified students by using genetic algorithms to predict their final grades.

Researchers used the data of university students in the e-learning environment called LON-CAPA. Four different classifiers were used in the study. The effective optimization of student classification in all three cases indicates the advantages of the usage of LON-CAPA data to predict the final grades of students based on their features extracted from the homework data. Kotsiantis & Pintelas (2005) predicted a student's marks (pass and fail classes) by use of regression techniques over Hellenic Open University data. In that study, six different classification algorithms were used.

That study concluded that M5rules was the most accurate regression algorithm that could be used for the construction of a software support tool. Furthermore, another advantage of M5rules, besides its superior performance, was its better comprehensibility.

Vandamme et al. (2006) made an attempt to classify students into three groups: 'low-risk' students having a high probability of succeeding; 'medium-risk' students who may be successful if the university takes appropriate measures; and 'high-risk' students with a high possibility of failing or dropping out.

In that study, artificial neural networks, decisions trees, and a linear discriminant analysis were used for classification purposes. According to the results of that study, linear discriminant analysis is the most effective method for classification.

Literature review shows that the EDM applications on e-learning environments are mainly aimed at conducting the automatic analysis of learner interaction and behavioral data via e-learning environments (Abdous & He, 2011; Chen et al., 2007; He, 2013; Jovanovica et al., 2012; Lazcorreta et al., 2008; Lee et. al., 2009; Romero & Ventura, 2007; Romero & Ventura, 2010; Romero et al., 2010; Zafra et al., 2010).

There are also many studies making an attempt to classify the exam performances of students (Falakmasir & Jafar, 2010; Fausett & Elwasif, 1994; Kotsiantis & Pintelas, 2005; Minaei-Bidgoli & Punch, 2003). Most of the studies have been conducted at university level. In the present study, the ability levels of the 11th grade students measured through computerized adaptive test were classified.

Based on the classification, the ability levels of students about permutation, combination, binomial expansion, and probability, which were the sub-topics of the probability unit, were evaluated. This is an authentic study in that the data acquired from high school students (probability unit ability level) were used, and the ability levels of students were tested through computerized adaptive test.

METHODOLOGY

In this study, a CAT system was developed for permutation, combination, binomial expansion, and probability, which were the sub-topics of the 11th grade mathematics course probability unit. An exam was conducted for the probability unit through the CAT system developed. The exam was conducted in five sessions (i.e. permutation test, combination test, binomial expansion test, probability test, and end-of-unit test). The ability levels acquired from the exam were analyzed and classified through hierarchical clustering method via SPSS 16.0 packages.

Procedure

In this study, the 11th grade students learnt the probability unit through a method different from traditional approach. Students were taught permutation, combination, binomial expansion, and probability, which were the sub-topics of the probability unit, in the individualized mathematics learning environment called UZWEBMAT (XXX, 2013). At the end of the teaching of each sub-topic, students were put into exam through the CAT module addressing the related subject.

In addition, students were subjected to an end-of-unit test consisting of questions about the entire probability unit after they had undergone permutation, combination, binomial expansion, and probability tests. In this way, students participated in 5 exam sessions.

Sample

An exam was conducted in school environment through the testing system developed, and the acquired data were evaluated. This exam was carried out in a high school located in Trabzon, Turkey during the 2011-2012 academic year fall semester. A total of 84 eleventh grade students participated in the exam. The names of students were not used. Instead, students were coded as follows: Std1, Std2, ... , Std84.

Data Collection Tool

System records were used for data collection. System records contained the ability levels and score details of each student concerning all exams.

FINDINGS

An exam was conducted about each sub-topic of the probability unit by using the CAT system. Hierarchical clustering method was used for clustering the ability levels obtained from each exam of 84 students participating in the study. The similarities between ability levels were demonstrated with dendrograms. The ability level clustering findings concerning permutation test, combination test, binomial expansion test, probability test, and end-of-unit test, are presented below. Firstly, clustering analysis concerning the ability levels obtained from permutation test is presented.

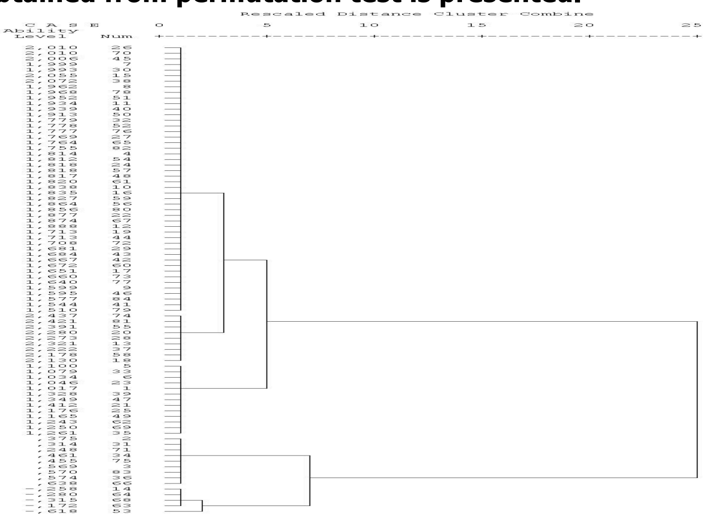

Figure: 1
Dendrograms regarding permutation test ability levels

Figure: 1 demonstrates the dendrograms regarding permutation test ability levels.

Based on the examination of figure 1, it is seen that the permutation test ability levels of students gather in two main clusters: A and B.

The examination of dendrograms in figure 1 reveals that cluster A is made up of two subsets: A1 and A2.

In addition, subset A1 consists of two further subsets: A11 and A12. Similarly, cluster B consist of two further subsets: B1 and B2. Subset B2 has two further subsets: B21 and B22.

Details about the ability level values covered by the clusters A and B are provided below:

Cluster A

A1

A11: from 2.01 to 1.51
A12: from 2.13 to 2.43

A2: from 1.261 to 1.017

Cluster B

B1: from 0.638 to 0.314

B2

B21: from -0.172 to -0.315
B22: -0.618

Figure: 2 presents the dendrograms regarding combination test ability levels. Based on the examination of figure 2, it is seen that the combination test ability levels of students gather in two main clusters: C and D.

The examination of dendrograms in figure 2 reveals that cluster C is made up of two subsets: C1 and C2. In addition, subset C1 consists of two further subsets named C11 and C12 and subset C2 consists of two further subsets: C21 and C22.

Similarly, cluster D consist of two further subsets: D1 and D2. Details about the ability level values covered by the clusters C and D are provided below:

Cluster C

C1

C11: from 2.041 to 1.442
C12: from 2.481 to 2.078

C2
C21: from 1.376 to 1.051
C22: from 0.938 to 0.554

Cluster D

D1: from -0.179 to -0.427
D2: from 0.342 to -0.089

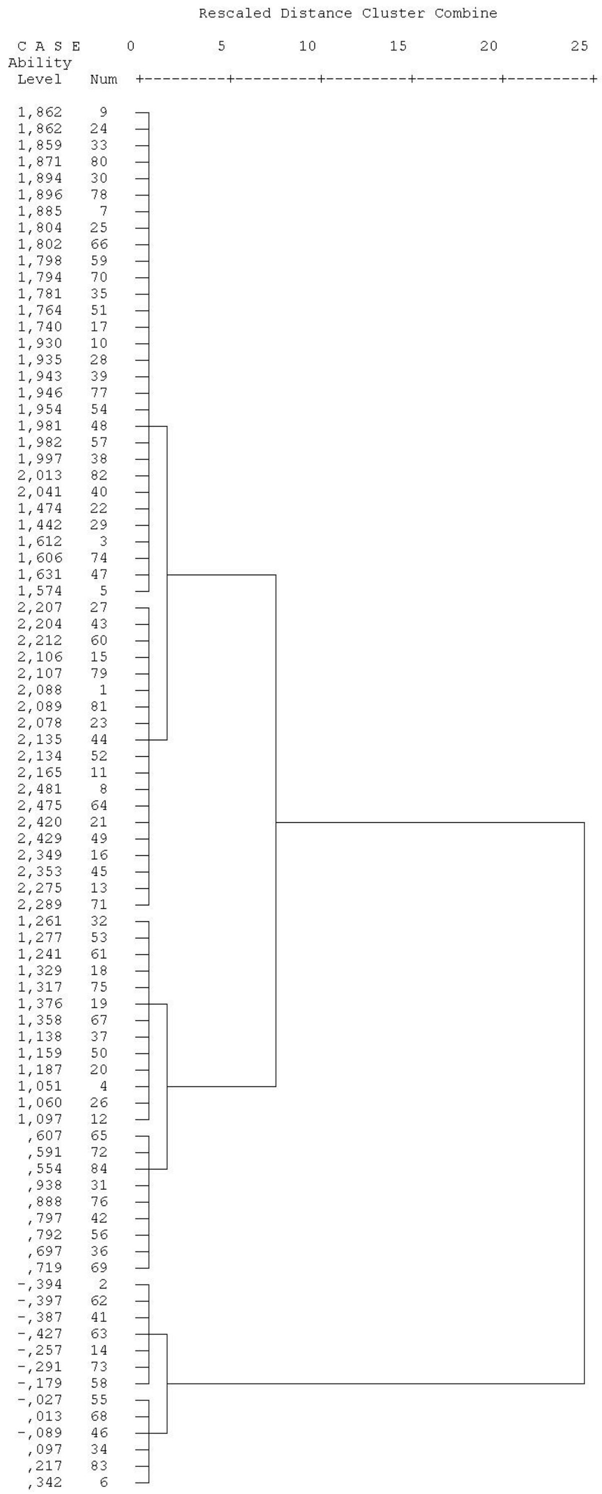

Figure: 2
Dendrograms regarding combination test ability levels

Figure: 3 presents the dendrograms regarding binomial expansion test ability levels. Based on the examination of figure 3, it is seen that the binomial expansion test ability levels of students gather in two main clusters: E and F.

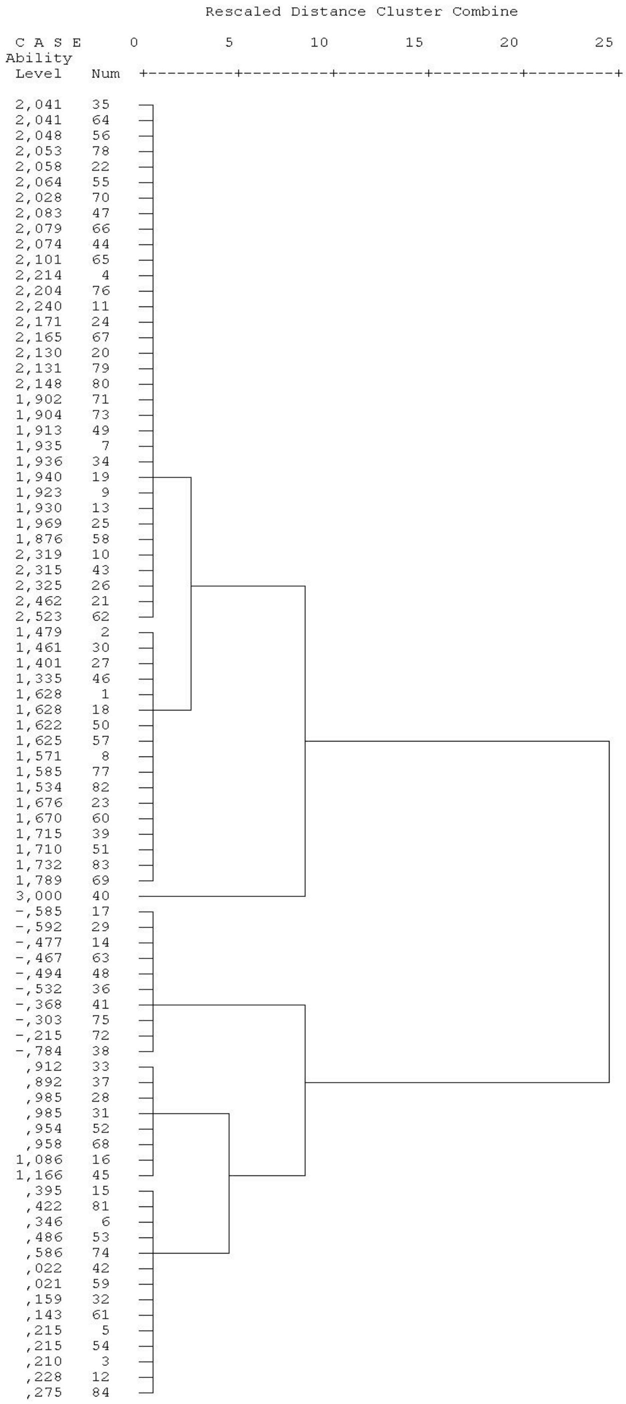

Figure: 3
Dendrograms regarding binomial expansion test ability levels

The examination of dendrograms in figure 3 reveals that cluster E is made up of three subsets: E1, E2, and E3. Similarly, cluster F consist of two further subsets:

F1 and F2. In addition, subset F1 consists of two further subsets named F21 and F22. Details about the ability level values covered by the clusters E and F are provided below:

Cluster E

E1

E11: from 2.523 to 1.876
E12: from 1.789 to 1.335

E2: 3

Cluster F

F1: from -0.215 to -0.592

F2

F21: from 1.166 to 0.892
F22: from 0.586 to 0.021

Figure: 4 presents the dendrograms regarding probability test ability levels. Based on the examination of figure 4, it is seen that the probability test ability levels of students gather in two main clusters: G and H.

The examination of dendrograms in figure 4 reveals that cluster G is made up of two subsets: G1 and G2. Similarly, cluster H consists of two further subsets: H1 and H2. In addition, subset H1 consists of three further subsets named H11, H12, and H13, and subset H12 consists of two further subsets: H21 and H22. Details about the ability level values covered by the clusters G and H are provided below:

Cluster G

G1: from 0.027 to -0.301
G2: from -0.411 to -0.71

Cluster H

H1

H11: from 1.779 to 1.591
H12: from 2.436 to 2.335
H13: from 2.239 to 1.838

H2

H21: from 1.452 to 0.863
H22: from 0.682 to 0.164

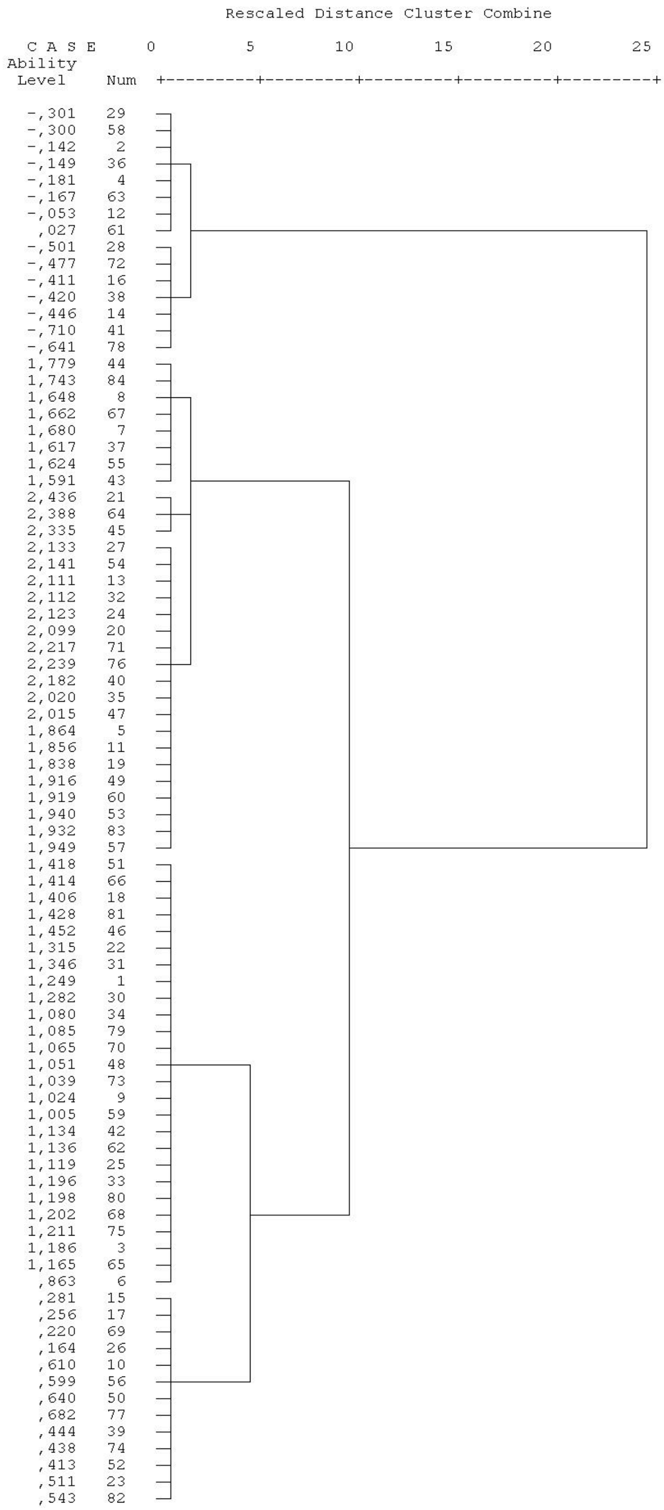

Figure: 4
Dendrograms regarding probability test ability levels

Figure: 5 presents the dendrograms end-of-unit probability test ability levels. Based on the examination of figure 5, it is seen that the end-of-unit test ability levels of students gather in two main clusters: K and L.

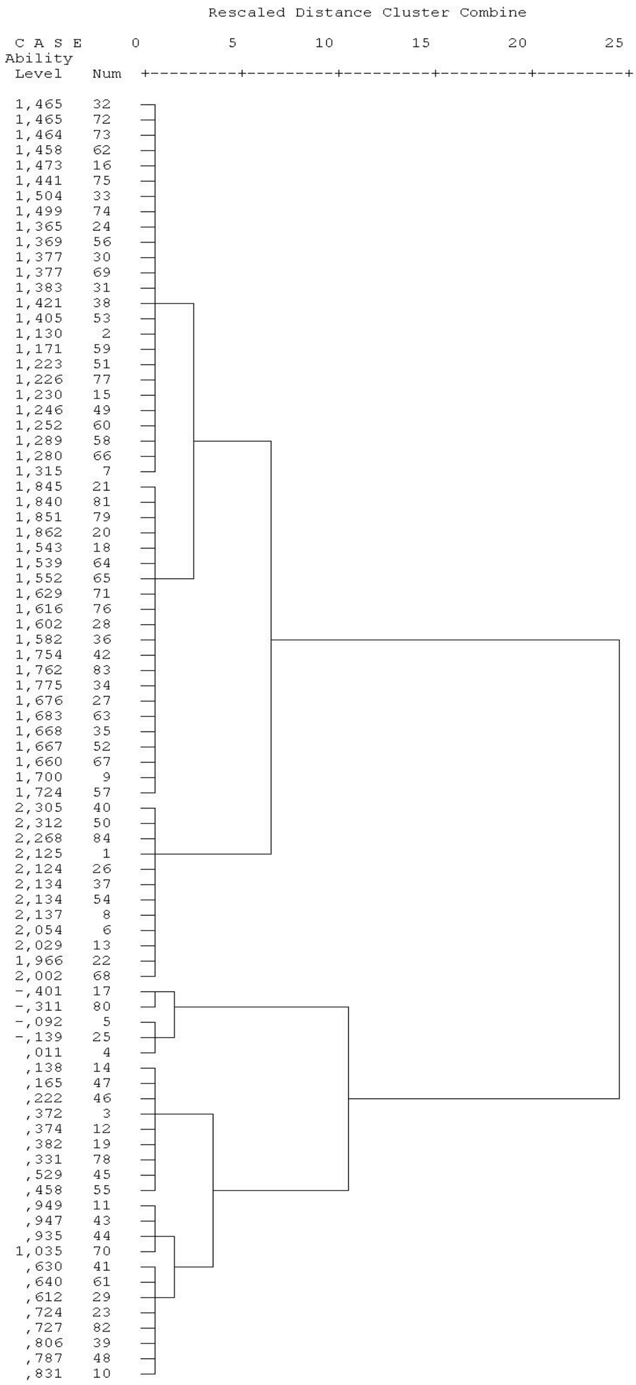

Figure: 5
Dendrograms regarding end-of-unit test ability levels

The examination of dendrograms in figure 5 reveals that cluster K is made up of two subsets: K1 and K2. In addition, subset K1 consists of two further subsets named K11 and K12. Similarly, cluster L consists of two further subsets: L1 and L2. Moreover, subset L1 consists of two further subsets named L11 and L12, and subset L2 is made up of two further subsets: L21 and L22. Furthermore, subset L22 consists of two further subsets: L221 and L222. Details about the ability level values covered by the clusters K and L are provided below:

Cluster K

K1

K11: from 1.504 to 1.171
K12: from 1.862 to 1.539

K2: from 2.312 to 1.966

Cluster L

L1

L11: from -0.311 to -0.401
L12: from 0.011 to -0.139

L2

L21: from 0.529 to 0.138

L22

L221: from 1.035 to 0.935
L222: from 0.63 to 0.831

CONCLUSIONS

This study classified the ability levels of the eleventh grade Turkish students concerning the sub-topics of the probability unit. The probability unit ability levels of students were obtained through CAT application integrated into UZWEBMAT environment.

The findings of the present study can be summarized as follows: the ability levels of students gather in two main clusters for the permutation test. While the first main cluster contains the values from 2.43 to 1.017, the second main cluster covers the values between 0.638 and -0.618.

Based on the examination of the ranges of these two main clusters, it is seen that the ability levels of students gather in two main clusters: medium ability level and advanced ability level according to the ability level variation scale (-3 to +3). The knowledge levels of students gather in two main clusters for the combination test.

While the first main cluster contains the values between 2.48 and 0.55, the second main cluster covers the values between 0.34 and -0.427. Based on the examination of the ranges of these two main clusters, it is seen that the combination test ability levels of students gather in two main clusters: medium ability level and advanced ability level.

The ability levels of students gather in two main clusters for the binomial expansion test. While the first main cluster contains the values from 3 to -1.335, the second main cluster covers the values between 1.116 and 0.021. Based on the examination of the ranges of these two main clusters, it is seen that the ability levels of students cluster above the medium ability level and at advanced ability level. In addition, the examination of the first main cluster shows that this cluster is divided into two within itself. While the first one of these subsets takes a value between 2.523 and 1.335, the second subset contains only one element (3).

This is because there is no other element between 3 and 2.523. It is seen that the probability test ability levels of students gather in two main clusters. While the first main cluster contains the values from 0.027 to -0.71, the second main cluster covers the values between 2.436 and 0.164. Based on the examination of the ability level ranges of these two main clusters, it is seen that the probability test ability levels of students gather in two clusters: medium ability level and advanced ability level. Finally, the end-of-unit test ability levels of students gather in two main clusters.

While the first main cluster contains the values from 2.312 to 1.504, the second main cluster covers the values between 1.035 and -0.401. Based on the examination of the ability level ranges of these two main clusters, it is seen that the end-of-unit test ability levels of students gather in two clusters: medium ability level and advanced ability level.

In conclusion, the probability unit ability levels of students gather in two clusters: medium and advanced knowledge levels.

BIODATA and CONTACT ADDRESSES of AUTHOR

Ozcan OZYURT was born in Trabzon, Turkey in 1978. He received the B.Sc. and M.Sc. degrees in Computer Engineering from Karadeniz Technical University (KTU) in 1996 and 2000, respectively. He completed his doctoral studies (Ph.D.) on adaptive and educational hypermedia in secondary mathematics education at Karadeniz Technical University, 2013. Now, he has been working as a full-time faculty member, Assist. Prof. Dr., in the Software Engineering Department, Technology Faculty, Karadeniz Technical University, Trabzon, Turkey. His major research interests are in the use of artificial intelligence in education, adaptive and intelligent tutoring system, distance education software, e-learning and mathematics education.

Assist. Prof. Dr. Özcan ÖZYURT
Karadeniz Technical University, OF Technology Faculty,
Software Engineering Department,
Of, Trabzon, TURKEY,
Phone: (0462) 7717250
Emails: oozyurt@ktu.edu.tr or oozyurt61@gmail.com

REFERENCES

Abdous, M., & He, W. (2011). Using text mining to uncover students' technology related problems in live video streaming. *British Journal of Educational Technology*, 42(1), 40–49.

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching & Learning*, 4(2), 1–9.

Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–17.

Chang, L. (2006). Applying data mining to predict college admissions yield: A case study. *New Directions for Institutional Research*, 2006(131), 53–68.

Chen, S.Y., & Liu, X. (2011): Mining students' learning patterns and performance in Web-based instruction: a cognitive style approach. *Interactive Learning Environments*, 19(2), 179-192.

Chen, C., Hsieh, Y., & Hsu, S. (2007). Mining learner profile utilizing association rule for web-based learning diagnosis. *Expert Systems with Applications*, 33(1), 6-22.

Falakmasir, M.H, & Jafar, H. (2010). *Using educational data mining methods to study the impact of virtual classroom in e-learning*. Paper presented at the Proceedings of the 3rd International Conference on Educational Data Mining, Pittsburgh, PA, USA.

Fausett, L., & Elwasif, W. (1994). *Predicting performance from test scores using backpropagation and counterpropagation*. In WCCI'94: IEEE world congress on computational intelligence. Washington, USA (pp. 3398–3402).

Garcia, E., Romero, C., Ventura, S., & de Castro, C. (2011). A collaborative educational association rule mining tool. *Internet and Higher Education*, 14(2), 77–88.

He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, 29(2013), 90–102.

Jovanovica, M., Vukicevica, M., Milovanovica, M., & Minovica, M. (2012). Using data mining on student behavior and cognitive style data for improving e-learning systems: A case study. *International Journal of Computational Intelligence Systems*, 5(3), 597–610.

Klosgen, W., & Zytkow, J. (2002). Handbook of data mining and knowledge discovery. New York: Oxford University Press.

Kotsiantis, S., & Pintelas, P. (2005). *Predicting students marks in hellenic open university*. In ICALT'05: The fifth international conference on advanced learning technologies. Kaohsiung, Taiwan (pp. 664–668).

Kotsiantis, S., Patriarcheas, K., Xenos, M. (2010). A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education. *Knowledge-Based Systems*, 23(2010), 529–535.

Kreitzber, C. B., Stocking, M. L. & Swanson, L. (1978). Computerized adaptive testing: Principles and directions. *Computers & Education*, 2, 319-329.

Lazcorreta, E., Botella, F., & Fernández-Caballero, A. (2008). Towards personalized recommendation by two-step modified apriori data mining algorithm. *Expert Systems with Applications*, 35(3), 1422–1429.

Lee, M.W., Chen, S.Y., Chrysostomou, K., & Liu, X. (2009). Mining students' behavior in web-based learning programs. *Expert Systems with Applications*, 36(2009) 3459-3464.

Lee, Y-J. (2012). Developing an efficient computational method that estimates the ability of students in a Web-based learning environment. *Computers & Education*, 58(2012), 579–589.

Levy, S.T., Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education*, 56(2011), 556–573.

Liao, S., Chu, P., & Hsiao, P. Y. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications*, 39(2012), 11303–11311.

Mostow, J., & Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering*, 12(2), 195–208.

Nandeshwar, A, Menzies, T., & Nelson, A. (2011). Learning patterns of university student retention. *Expert Systems with Applications*, 38(2011), 14984–14996.

Ozyurt, O., Ozyurt, H., & Baki, A. (2013). Design and development of an innovative individualized adaptive and intelligent e-learning system for teaching-learning of probability unit: Details of UZWEBMAT. *Expert Systems with Applications*, 40(8), 2914-2940.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(2007), 135–146.

Pal, S. (2012). Mining educational data to reduce dropout rates of engineering students. *International Journal of Information Engineering and Electronic Business*, 2(1), 1–7.

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 40(6), 601–618.

Romero, C., Ventura S., & Garcia E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(2008), 368–384.

Romero, C., Espejo, P.G., Zafra, A., Romero, J.R., & Ventura, S. (2010). Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education*. doi: 10.1002/cae.20456.

Tsantis, L., & Castellani, J. (2001). Enhancing learning environments through solution-based knowledge discovery tools. *Journal of Special Education Technology*, 16(4), 1–35.

Vandamme, J.P., Meskens, N. & Superby, J.F. (2007). Predicting academic performance by data mining methods. *Education Economics*, 15(4), 405-419.

Wang, Y-H., & Liao, H-C. (2011). Data mining for adaptive learning in a TESL-based e-learning system. *Expert Systems with Applications*, 38(2011), 6480–6485.

Weiss, D. J. (1985). Adaptive testing by computer. *Journal of Consulting and Clinical Psychology*, 53(6), 774-789.

Zafra, A., Romero, C., & Ventura S. (2011). Multiple instance learning for classifying students in learning management systems. *Expert Systems with Applications*, 38(2011), 15020, 15031.