Elementary Education Online, 14(3), 815-824, 2015.
İlköğretim Online, 14(3), 815-824, 2015. [Online]: http://ilkogretim-online.org.tr
http://dx.doi.org/10.17051/io.2015.03160

Modeling Students' Academic Performance Based on Their Interactions in an Online Learning Environment[i]

Gökhan AKÇAPINAR[*], Arif ALTUN[**], Petek AŞKAR[***]

ABSTRACT. The aim of this study is to model students' academic performance based on their interactions in an online learning environment. The dataset includes 10 input attributes extracted from students' learning interaction data. As an output (class) variable, the final grades obtained from their Computer Hardware course were used. The modeling performance of three different classification algorithms were tested (naïve Bayes classifier, classification tree and CN2 rules) on the dataset. All analyses were performed using the Orange data mining tool, and the models were evaluated using ten-fold cross-validation. The results of analysis were presented as a confusion matrix, a decision tree, and if-then rules. The predictive performance of the algorithms was also tested and compared using the classification accuracy (CA), and area under the ROC Curve (AUC) metrics. The experimental results indicate that the naïve Bayes algorithm outperforms other classification algorithms when compared using the CA and AUC metrics. The naïve Bayes algorithm correctly classified 75.4% of the students according to their grade for the course (Fail, Pass, and Good). The classification model also accurately predicted 81.5% of the students who failed, and 91.8% of the students who passed the course. On the other hand, the classification tree and the CN2 algorithms generated models which can be used with confidence in decision making processes by non-expert data mining users.
Key Words: academic performance modeling, final grade prediction, classification, educational data mining, learning analytics.

INTRODUCTION

Online learning environments have many tools for supporting teaching and learning. They enable educators to share course content, prepare assignments and tests as well as to engage in discussions (Cristobal Romero, Espejo, Zafra, Romero, & Ventura, 2013). They also support collaborative learning with tools such as wikis, forums, chats and so forth (Moreno, Gonzalez, Castilla, Gonzalez, & Sigut, 2007). In addition to these advantages, all kinds of online learning environments produce significant amount of interaction data regarding students' learning processes (Greller & Drachsler, 2012; Koedinger, Cunningham, Skogsholm, & Leber, 2008). Although these data are recorded automatically in databases, using them for the purpose of improving education is restricted to simple statistics and graphics (Ali, Asadi, Gašević, Jovanović, & Hatala, 2013).

Educational Data Mining (EDM) and Learning Analytics (LA) are emerging fields related to the usage of the data collected in educational environments to improve education (Siemens & Baker, 2012). In recent years, analyzing educational data by using Data Mining (DM) techniques (e.g., classification, clustering, association rules) have been used in many tasks such as predicting students' off-task (Baker, 2007) and gaming (Beal, Qu, & Lee, 2008) behaviors, predicting students' level of disorientation (Akçapınar, Cosgun, & Altun, 2011), grouping similar students (Kardan & Conati, 2011), recommending e-learning materials (Jie, 2004) and offering e-learning courses to students (Alfredo, Félix, & Àngela, 2010).

This study aimed to model students' academic performance based on features extracted from their online learning environment usage data. Students' final grade obtained in the Computer Hardware course (Fail, Pass, and Good) was considered as an indicator of the students' academic performance.

Academic Performance Modeling

Modeling students' academic performance is one of the popular applications of DM in educational settings (Bousbia & Belamri, 2014; C. Romero & Ventura, 2010). A recent review study

[i] This study was produced from the PhD thesis of the first author.

[*] Dr., Hacettepe University, Faculty of Education, Department of Computer Education & Instructional Technology, Ankara, Turkey. gokhana@hacettepe.edu.tr
[**] Prof. Dr., Hacettepe University, Faculty of Education, Department of Computer Education & Instructional Technology, Ankara, Turkey. altunar@hacettepe.edu.tr
[***] Prof. Dr., Hacettepe University, Faculty of Education, Department of Computer Education & Instructional Technology, Ankara, Turkey. paskar@ hacettepe.edu.tr

has shown that student performance modeling has increased its popularity in the last three years (Peña-Ayala, 2014). The main goal of these studies is to predict how successfully the learner is or will be able to complete a given task and/or to achieve a specific learning goal (Peña-Ayala, 2014).

There are several studies of modeling students' academic performance. For example, Osmanbegović and Suljić (2012) posited that students' demographics and precollege academic success data gathered with a self-report survey could be used to predict their success in a course. In another study, Akçapınar, Çoşgun, and Altun (2013) conducted a prediction study for students' final grades based on their interaction with a wiki environment. Their feature set included students' session and navigation metrics as well as wiki based metrics such as edit count and word count. Their results showed that the support vector machine algorithm can be trained to predict final grades with an accuracy of 67.1%. Lopez, Luna, Romero, and Ventura (2012) applied classification via clustering to predict the final grades of students in a course based on students' participation in Moodle forum. Their feature set included ten variables related to students' forum usage. They found that the EM clustering algorithm yields results similar to those of the best classification algorithms. McCuaig and Baldwin (2012) asserted that the source log data produced by conventional LMS could be mined to predict the students' success or failure without requiring the results of formal assessments. Márquez-Vera, Cano, Romero, and Ventura (2013) also conducted a study to predict which students might fail a course using students' online performances.

Academic performance modeling based on features that has been extracted from students' interaction logs is especially important for the early prediction of students' drop-out and their probability of failure in an online course. Moreover, these models can be used to classify students automatically or to make an automatic adaptation based on their activity level in an adaptive learning environment. Teachers can use these models to monitor their students' learning progress from predesigned and/or interactive dashboards instantly.

METHOD

Data mining, which is used successfully in many fields to discover hidden patterns and relationships in data, was the research method for this study. The procedures in the data mining process are presented in Figure 1.

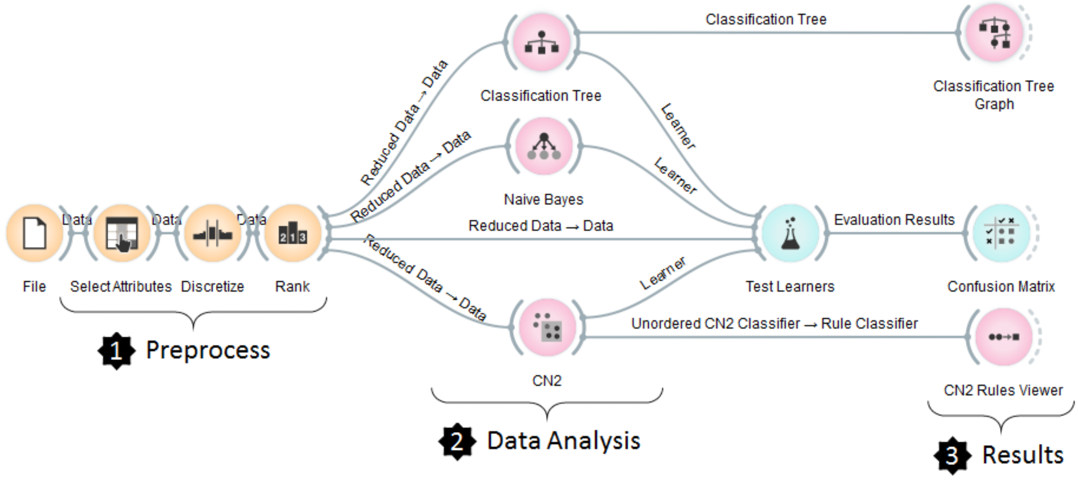

Figure 1. Data mining process

Participants

In the 2013-2014 academic year, 76 undergraduates (41 female, 35 male) in the Computer Hardware course in the Computer Education and Instructional Technology Department participated in this study. Their ages ranged from 19 to 23.

Procedure

In addition to face to face classes, students performed activities in an online learning environment designed by the first author in this study to assist their learning. These activities included writing reflections about the concepts which they learned in the course, reading sources related to the course, commenting on and assessing posts written by other students, asking questions in discussions, writing answers to questions, assessing questions and answers which were written by others, keeping track of announcements and using the course resources. In addition to data about these activities, students' navigation and session logs are stored in a database within the system.

Dataset

The dataset used in the study was obtained from the database in an online learning environment. The raw data contains fourteen weeks of usage by 76 students with 3,803 logins, 4,130 posts, 3,937 tags, and more than 100,000 page views. Before applying any data mining algorithms, the raw data was analyzed and the features—representing students' learning behaviors—were extracted automatically by a feature extraction tool. The list of the features is shown in Table 1. The last one is the categorical final grade which reflects students' performance and was added manually to the analysis table.

Table 1. Description of all features

No	*Attribute*	*Description*
1	n_Login	login count
2	d_Usage	total time in minutes. spent in the environment
3	n_Post	post count
4	n_Tag	number of tags used in posts
5	n_PostNav	number of navigations to posts written by other students
6	n_PostAss	number of assessment of posts written by other students
7	n_Answer	number of written responses to questions in the discussion section
8	n_DissNav	number of navigations to the discussion section
9	n_AnswerNav	number of navigations to the questions and answers
10	n_QuestionAss	number of assessments of questions in the discussion section
11	f_Grade	final grades

Preprocessing

All continuous features, except for final grades, were discretized to provide a more comprehensible view of the data with the help of the entropy-MDL algorithm (Márquez-Vera et al., 2013). Entropy-MDL is a class-aware discretization introduced by Fayyad and Irani (1992) that uses MDL and entropy to find the best cut-off points. The final grades were manually discretized into three intervals (Fail, Pass, and Good) according to the course grading policy of the university (scores between 0 – 49 coded as Fail, 50 – 69 coded as Pass, and 70 – 100 coded as Good). The distribution of the students in these categories are: 35.53% Fail, 36.84% Pass and 27.63% Good.

In the second step of preprocessing, the feature selection algorithm was applied to the data using the Orange Rank widget. The rank widget scores the features based on selected scoring techniques, e.g., gini gain, information gain, linear SVM weights (Orange, 2014). The feature selection process can be used to reduce the number of features (also known as important features) without losing reliability in classification (Lopez et al., 2012). The linear SVM weights were used to rank features as detailed in (Chang & Lin, 2008). The result of the algorithm is shown in Table 2. The prediction performance of the algorithms was tested before and after applied feature selection.

Table 2. The SVM weight of features

No	*Attribute*	*SVM weights*
1	n_Login	2.382
2	n_Post	1.565
3	n_QuestionAss	1.411
4	n_Answer	0.719
5	n_Tag	0.556
6	d_Usage	0.496
7	n_AnswerNav	0.399
8	n_PostNav	0.248
9	n_DissNav	0.080
10	n_PostAss	0.065

Data Analysis

Choosing the most suitable algorithm for a new dataset is important task since there is no single classifier that yields the best results on all datasets (Osmanbegović & Suljić, 2012; Cristobal Romero, Espejo, Zafra, Romero, & Ventura, 2010). Therefore, as a first step, the predictive performance and intelligibility of the models generated by three different algorithms were compared. The algorithms were the naïve Bayes classifier (NB), the classification tree (CT), and CN2 rules. The following section gives a brief introduction to these algorithms.

Naïve Bayes: A naïve Bayes classifier is a simple probabilistic classifier based on Bayes' theorem (Hongbo, Yizhou, Yi, & Jiawei, 2014). Bayesian classifiers are popular classification algorithms due to their simplicity, computational efficiency and very good performance for real-world problems (Kabakchieva, 2013).

Decision Tree: Decision trees are hierarchical representations of data (Charu, 2014). While the top node is called the root, lower nodes are called leaves, and each of them represents a class. The variables and the splitting points where the data will be divided in each step are determined by the decision tree algorithm being used. While the C4.5 algorithm (Quinlan, 1993) uses the information gain ratio, the CART algorithm (Breiman, Friedman, Olshen, & Stone, 1984) uses the Gini index as a splitting criteria. In this study, the Orange software's classification tree algorithm was used with the Gini index parameter.

CN2 Rules: Rule based algorithms are similar to decision trees. They differ from decision trees since there is no strict hierarchical partitioning. For example, a path in a decision tree can be thought as a rule in a rule based algorithm (Charu, 2014). Rules are described as if-then clauses. The IF clause contains a combination of conditions for the predicting attributes. The THEN clause contains the predicted value for the class (Cristobal Romero et al., 2013). Numerous methods, such as classification based on associations (Liu, Hsu, & Ma, 1998) and CN2 (Clark & Niblett, 1989), have been proposed in the literature. They use a variety of rule induction methods, based on different ways of mining and prioritizing rules.

Two sets of experiments were carried out to evaluate the effects of feature selection technique on the predictive performance of models. The first used all the features, and the second used the five best ranked features (n_Login, n_Post, n_QuestionAss, n_Answer and n_Tag according to their SVM weights given in Table 2). In addition to the models generated by these algorithms, their predictive performances are presented in the results section. To evaluate the predictive performance of the models, two metrics were computed: classification accuracy (CA), and area under the ROC curve (AUC). All the analyses were performed using the Orange data mining tool (Demšar et al., 2013) with default parameters, and the models obtained were evaluated using ten-fold cross-validation.

RESULTS

The results of the first and second experiments are shown in Table 3. According to these results, the naïve Bayes algorithm outperforms the others in the first and second experiments. On the other hand, it can be seen from the table that other algorithms also yielded improved classification accuracy in the second experiment.

Table 3. Experimental results

Algorithm	***1st Experiment***		***2nd Experiment***	
	CA	***AUC***	***CA***	***AUC***
Classification Tree	0.711	0.815	0.727	0.824
CN2 rules	0.555	0.763	0.711	0.834
Naïve Bayes	**0.721**	**0.883**	**0.754**	**0.871**

Naïve Bayes Algorithm Results

Although it is called naïve because of its assumptions (Osmanbegović & Suljić, 2012), the models generated by Bayesian networks, including the naïve Bayes classifier, tend to be difficult to understand for non-expert users (Xing, Guo, Petakovic, & Goggins, 2014). Therefore, we presented the naïve Bayes classifier results as a confusion matrix. The confusion matrix obtained by cross-validation is a fair indicator of the predictive performance of the algorithm on independent test samples (Enot et al., 2008). Table 4 shows the confusion matrix of the naïve Bayes algorithm obtained by cross-validation.

Table 4. Naïve Bayes Confusion Matrix*

	FAIL	PASS	GOOD
FAIL	22	5	0
PASS	3	19	6
GOOD	1	4	16

* Columns represent predictions, rows represent true classes

As indicated in Table 4, although there were some misclassification between fail and pass (n = 8) or pass and good classes (n = 10), only one student was classified as a fail which should have been good.

Decision Tree Algorithm Results

Figure 2 presents a sample decision tree model generated by the classification tree algorithm. When we analyze the decision tree in the figure, we can see that classifications of students can be predicted by two variables: n_QuestionAss and n_Answer (see Table 1 for description of features).

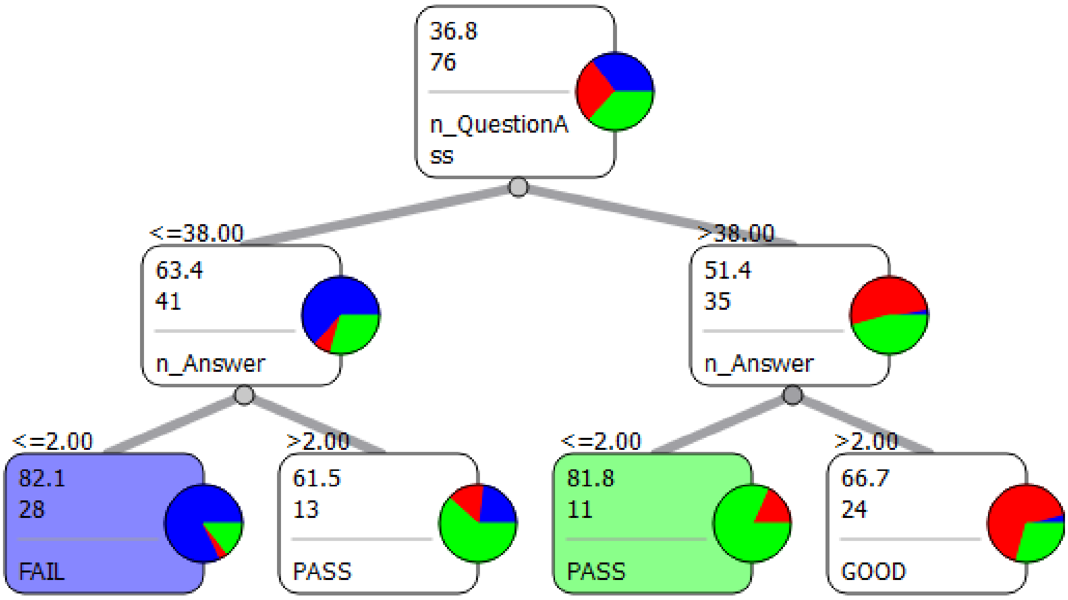

Figure 2. Sample tree from the decision tree algorithm

When we analyze the classification of the students by academic performance from top to bottom, we obtain the following results. Students with a total assessment count of more than 38 and, on the discussion board, answers more than 2 questions will be classified as good with a probability of 66.7% (n = 24). On the other hand, those with an assessment count less than 38 and answers less than 2 questions on the discussion board will be classified as fail with a probability of 82.1% (n = 28).

CN2 Algorithm Results

The series of rules in Figure 3 generated with CN2 algorithm show that two rules can be used to classify students as fail or pass/good. Rule 1 indicates that students with login counts of less than 17 and their post count is less than 12 will be classified as fail (n = 13). Rule 4 indicates that students having a question assessment count of more than 38, answers more than 2 questions on the discussion board, and a tag usage count of more than 11 will be classified as good or pass (n = 20).

	Rule quality	Coverage	Predicted class	Distribution	Rule
1	0.93	13.00	FAIL	<13.0,0.0,0.0>	IF n_Login=<=17.00 AND n_Post=<=12.00 THEN f_Grade=FAIL
2	0.75	6.00	FAIL	<5.0,0.0,1.0>	IF n_QuestionAss=<=38.00 AND n_Answer=<=2.00 AND n_Post=<=12.00 THEN f_Grade=FAIL
3	0.58	10.00	FAIL	<6.0,1.0,3.0>	IF n_QuestionAss=<=38.00 AND n_Answer=<=2.00 AND n_Post=>12.00 THEN f_Grade=FAIL
4	0.73	20.00	GOOD	<0.0,15.0,5.0>	IF n_QuestionAss=>38.00 AND n_Answer=>2.00 AND n_Tag=>11.00 THEN f_Grade=GOOD
5	0.82	9.00	PASS	<0.0,1.0,8.0>	IF n_Tag=>11.00 AND n_Answer=<=2.00 AND n_QuestionAss=>38.00 THEN f_Grade=PASS
6	0.64	12.00	PASS	<2.0,2.0,8.0>	IF n_Answer=>2.00 AND n_QuestionAss=<=38.00 AND n_Login=>17.00 THEN f_Grade=PASS

Figure 3. Sample rules from CN2 algorithm

The first rule indicates that student with a low number of logins and who write low numbers of posts will probably classified as fail. The second rule shows that students with a high number of question assessments, answers more questions on the discussion board and use more tags will be classified as pass or good.

DISCUSSION AND CONCLUSIONS

This paper compares three different classification algorithms that model students' academic performance. Along with the widely used features (login count, session duration, answer count in discussion board etc.), we added some new features such as tag usage count, assessment count, and navigation count which reflect students' learning behavior in the online learning environment. We carried out two experiments to test the effects of feature selection algorithm on classification performance. While other algorithms improved their classification performance in the second experiment, the highest classification accuracy was achieved by the naïve Bayes algorithm, which correctly classified 75.4% of the students by course grade.

The classification model accurately predicted 22 of the 27 students who failed (81.5%) and 45 of the 49 students who passed (91.8%). However, as mentioned by Osmanbegović and Suljić (2012), in the educational context not only classification accuracy, but also the ease of learning and the user friendly characteristics of the results are important to integrating these algorithms into learning environments. Bayesian networks including the naïve Bayes classifier are able to attain very good accuracy rates, but very difficult to understand for non-expert end users (e.g., teachers, students) (Xing et al., 2014). Therefore, they can be used if classification performance is more important than model interpretation (Dreiseitl & Ohno-Machado, 2002), for example when identifying at-risk students before the end of the semester. On the other hand, tree-based or rule-based algorithms generates models which are easy to understand for non-expert users (Osmanbegović & Suljić, 2012). They can be used to understand relationships between students' performances and underline the factors that affect their performances (Dreiseitl & Ohno-Machado, 2002). For example, teachers can use the decision tree and if-then rules generated here to make decisions about their students or give them feedback to improve their performance and reduce failure rates. Therefore, the advantages and disadvantages of the intelligibility of model and its predictive performance must be taken into account when choosing the best algorithm (Xing et al., 2014).

This study shows that students' login counts on to the online learning environment, login durations, participation in online discussions, writing reflections about concepts and tagging these reflections affect students' academic performance. In further research, these models can be used to determine student performance and to arrange early interventions before students fail or drop out. These systems are also referred to as early warning systems in the EDM and LA literature. Instructional designers can design online learning environments as well as instant dashboards based on these models. They can also arrange their instructional designs using these variables and enhance their students' participation.

REFERENCES

Akçapınar, G., Cosgun, E., & Altun, A. (2011). *Prediction of Perceived Disorientation in Online Learning Environment with Random Forest Regression.* Paper presented at the Proceedings of the 4th International Conference on Educational Data Mining., July 6-8, Eindhoven, Netherlands.

Akçapınar, G., Çoşgun, E., & Altun, A. (2013, October 17th - 18th). *Mining Wiki Usage Data for Predicting Final Grades of Students.* Paper presented at the International Academic Conference on Education, Teaching and E-learning (IAC-ETeL 2013), Prague, Czech Republic.

Alfredo, V., Félix, C., & Àngela, N. (2010). Clustering Educational Data *Handbook of Educational Data Mining* (pp. 75-92): CRC Press.

Ali, L., Asadi, M., Gašević, D., Jovanović, J., & Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education, 62*(0), 130-148. doi: http://dx.doi.org/10.1016/j.compedu.2012.10.023

Baker, R. S. J. d. (2007). *Modeling and understanding students' off-task behavior in intelligent tutoring systems*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA.

Beal, C. R., Qu, L., & Lee, H. (2008). Mathematics motivation and achievement as predictors of high school students' guessing and help-seeking with instructional software. *Journal of Computer Assisted Learning, 24*(6), 507-514. doi: 10.1111/j.1365-2729.2008.00288.x

Bousbia, N., & Belamri, I. (2014). Which Contribution Does EDM Provide to Computer-Based Learning Environments? In A. Peña-Ayala (Ed.), *Educational Data Mining* (Vol. 524, pp. 3-28): Springer International Publishing.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth International Group.

Chang, Y.-W., & Lin, C.-J. (2008). *Feature Ranking Using Linear SVM.* Paper presented at the JMLR Workshop and Conference Proceedings: Causation and Prediction Challenge (WCCI 2008).

Charu, C. A. (2014). An Introduction to Data Classification *Data Classification* (pp. 1-36): Chapman and Hall/CRC.

Clark, P., & Niblett, T. (1989). The CN2 Induction Algorithm. *Machine Learning, 3*(4), 261-283. doi: 10.1023/A:1022641700528

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., . . . Zupan, B. (2013). Orange: data mining toolbox in python. *J. Mach. Learn. Res., 14*(1), 2349-2353.

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics, 35*(5–6), 352-359. doi: http://dx.doi.org/10.1016/S1532-0464(03)00034-0

Enot, D. P., Lin, W., Beckmann, M., Parker, D., Overy, D. P., & Draper, J. (2008). Preprocessing, classification modeling and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data. *Nat. Protocols, 3*(3), 446-470.

Fayyad, U. M., & Irani, K. B. (1992). *The attribute selection problem in decision tree generation*. Paper presented at the Proceedings of the tenth national conference on Artificial intelligence, San Jose, California.

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society, 15 %6*(3), 42-57.

Hongbo, D., Yizhou, S., Yi, C., & Jiawei, H. (2014). Probabilistic Models for Classification *Data Classification* (pp. 65-86): Chapman and Hall/CRC.

Jie, L. (2004). *A personalized e-learning material recommender system*.

Kabakchieva, D. (2013). Predicting student performance by using data mining methods for classification. *Cybernetics and Information Technologies, 13*(1), 61-72. doi: 10.2478/cait-2013-0006

Kardan, S., & Conati, C. (2011, 2011). *A Framework for Capturing Distinguishing User Interaction Behaviours in Novel Interfaces.* Paper presented at the The 4th International Conference on Educational Data Mining (EDM 2011).

Koedinger, K., Cunningham, K., Skogsholm, A., & Leber, B. (2008). *An open repository and analysis tools for fine-grained, longitudinal learner data.* Paper presented at the Educational Data Mining 2008: 1st International Conference on Educational Data Mining, Proceedings.

Liu, B., Hsu, W., & Ma., Y. (1998). *Integrating Classification and Association Rule Mining.* Paper presented at the ACM KDD Conference.

Lopez, M. I., Luna, J. M., Romero, C., & Ventura, S. (2012). *Classification via clustering for predicting final marks based on student participation in forums.* Paper presented at the 5th International Conference on Educational Data Mining, EDM 2012, Chania, Greece.

Márquez-Vera, C., Cano, A., Romero, C., & Ventura, S. (2013). Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. *Applied Intelligence, 38*(3), 315-330. doi: 10.1007/s10489-012-0374-8

McCuaig, J., & Baldwin, J. (2012, Jun 19-21, 2012). *Identifying Successful Learners from Interaction Behaviour.* Paper presented at the 5th International Conference on Educational Data Mining (EDM), Chania, Greece.

Moreno, L., Gonzalez, C., Castilla, I., Gonzalez, E., & Sigut, J. (2007). Applying a constructivist and collaborative methodological approach in engineering education. *Computers & Education, 49*(3), 891-915. doi: http://dx.doi.org/10.1016/j.compedu.2005.12.004

Orange. (2014). from http://docs.orange.biolab.si/widgets/rst/data/rank.html

Osmanbegović, E., & Suljić, M. (2012). Data Mining Approach for Predicting Student Performance. *Economic Review, 10*(1).

Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications, 41*(4, Part 1), 1432-1462. doi: http://dx.doi.org/10.1016/j.eswa.2013.08.042

Quinlan, J. R. (1993). *C4.5: programs for machine learning*: Morgan Kaufmann Publishers Inc.

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*, n/a-n/a. doi: 10.1002/cae.20456

Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., & Ventura, S. (2013). Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education, 21*(1), 135-146. doi: 10.1002/cae.20456

Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40*(6), 601-618. doi: 10.1109/TSMCC.2010.2053532

Siemens, G., & Baker, R. S. J. d. (2012). *Learning Analytics and Educational Data Mining: Towards Communication and Collaboration.* Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.

Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2014). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior*(0). doi: http://dx.doi.org/10.1016/j.chb.2014.09.034

Öğrencilerin Akademik Performanslarının Çevrimiçi Öğrenme Ortamındaki Etkileşim Verilerine Göre Modellenmesi[ii]

Gökhan AKÇAPINAR[*], Arif ALTUN [**], Petek AŞKAR[***]

ÖZ. Bu çalışmanın amacı çevrimiçi öğrenme ortamındaki etkileşim verilerine göre öğrencilerin Bilgisayar Donanımı dersine ilişkin akademik performanslarının modellenmesidir. Çalışmada kullanılan veri seti öğrencilerin çevrimiçi öğrenme ortamındaki log verilerinden elde edilen 10 adet değişkeni ve sınıf (tahmin) değişkeni olarak da öğrencilerin akademik performanslarının yansıması olan dönem sonu notlarını içermektedir. Yapılan analizlerde 3 farklı veri madenciliği algoritmasının (Naïve Bayes, Karar Ağacı ve CN2) sınıflama performansı karşılaştırılmıştır. Elde edilen modellerin tahmin performanslarının karşılaştırılması için Doğru Sınıflama Oranı (DSO) ve ROC Altında Kalan Alan (EAKA) metrikleri kullanılmıştır. Tüm analizler Orange veri madenciliği yazılımı ile gerçekleştirilmiştir ve elde edilen modellerin genelleştirilmesi için 10k çapraz geçerlilik yöntemi kullanılmıştır. Analiz sonuçları çapraz tablo, karar ağacı ve eğer-ise kurallar dizisi şeklinde sunulmuştur.

Anahtar Kelimeler: Akademik performans modelleme, tahmin, sınıflama, eğitsel veri madenciliği, dönem sonu not tahmini.

ÖZET

Amaç ve Önem: Bu çalışmanın amacı öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verilerini kullanarak Bilgisayar Donanımı dersine ilişkin akademik performanslarının veri madenciliği yaklaşımı ile modellenmesidir. Akademik performansın modellenmesi son yıllarda eğitsel veri madenciliği ve öğrenme analitiği araştırmacılarının ilgilendiği önemli bir konudur. Burada amaç, oluşturulan modellerle öğrencinin bir öğrenme etkinliği sonucunda sergilediği ya da sergileyeceği performansın tahmin edilmesidir. Bu tür modeller, öğrenci performansının tahmin edilmesi, dersi bırakma ya da başarısız olma ihtimali yüksek olan öğrencilerin erkenden belirlenmesi ve uyarlanabilir öğrenme ortamlarında otomatik uyarlamalar yapmak amacıyla kullanılabilecek bilgiler üretilmesi açısından önemlidir.

Araştırmada sınıf (tahmin) değişkeni olarak öğrencilerin Bilgisayar Donanımı dersinden dönem sonunda aldıkları puanlar (başarısız, orta ve başarılı) kullanılırken girdi değişkenleri olarak öğrencilerin çevrimiçi öğrenme ortamındaki etkileşimlerini yansıtan 10 adet değişken kullanılmıştır. Öğrencilerin akademik performanslarını modellemek amacıyla literatürde sıklıkla kullanılan üç farklı sınıflama algoritması seçilmiş (Naïve Bayes, Karar Ağacı ve CN2) ve sonuçları karşılaştırılmıştır.

Yöntem: Araştırmada yöntem olarak, birçok farklı alanda verideki gizli örüntü ve ilişkileri keşfetmek amacıyla kullanılan ve başarılı sonuçlar üreten veri madenciliği süreci izlenmiştir. Çalışmaya Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü öğrencilerinden 2013 – 2014 Güz döneminde Bilgisayar Donanımı dersine kayıtlı 76 öğrenci katılmıştır. Araştırmada kullanılan veriler, birinci yazar tarafından geliştirilen çevrimiçi öğrenme ortamından elde edilmiştir. Öğrenciler sistemi araştırmacılar tarafından yürütülen Bilgisayar Donanımı dersi kapsamında 14 hafta süresince kullanmışlardır.

Öğrencilerin bu sistemde gerçekleştirdikleri temel aktiviteler; ders kaynaklarını takip etme, derste öğrendikleri kavramlarla ilgili yansıma yazma, tartışmalara katılma, duyuruları ve bildirimleri takip etme şeklindedir. Sistemin veri tabanında tutulan öğrenme aktiviteleri ile ilgili verilerin analizi ile araştırmada kullanılan değişkenler üretilmiştir. Veri madenciliği çalışmalarında önemli bir adım olan ön işleme sürecinde elde edilen tüm veriler ilk olarak kesikli hale dönüştürülmüştür. Daha sonra Orange yazılımında yer alan özellik seçme aracı kullanılarak değişkenlerin önemlilik katsayıları belirlenmiştir. Analiz aşamasında ise seçilen üç algoritma kullanılarak iki farklı deney

---

[ii] Bu çalışma birinci yazarın doktora tezinden üretilmiştir.

[*] Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, Ankara, Türkiye. gokhana@hacettepe.edu.tr

[**] Prof. Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, Ankara, Türkiye. altunar@hacettepe.edu.tr

[***] Prof. Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, Ankara, Türkiye. paskar@ hacettepe.edu.tr

gerçekleştirilmiştir. Birinci deneyde tüm değişkenler kullanılırken ikincide sadece önem derecesine göre (özellik seçme işlemi ile belirlenen) seçilen ilk beş değişken kullanılmıştır ve sonuçları karşılaştırılmıştır.

Bulgular: Analiz sonuçları incelendiğinde her üç algoritmanın da daha az değişkenin kullanıldığı ikinci deneyde seçilen performans metrikleri açısından (DSO ve EAKA) sınıflama performanslarını artırdığı gözlemlenmiştir. En iyi sınıflama oranına ise Naïve Bayes algoritması ile ulaşılmıştır. Elde edilen modelin uzman olmayan kişiler tarafından anlaşılması zor olduğu için Naïve Bayes algoritmasına ilişkin sonuçlar sınıflama performansını gösteren çapraz tablo şeklinde verilmiştir (Tablo 4.). Çapraz tablo incelendiğinde modelin bir öğrenci dışında Başarısız ve Başarılı öğrencilileri doğru olarak sınıfladığı görülmektedir.

Karar Ağacı algoritması ile elde edilen karar ağacı incelendiğinde (Şekil 2.) öğrencilerin yer aldığı sınıfların iki değişkene göre tahmin edilebileceği görülmektedir. Gini indeksi kriterine göre seçilen bu değişkenlerin tartışma ortamındaki soruları değerlendirme sayıları ve sorulara cevap yazma sayıları olduğu görülmektedir. Öğrencileri başarılarına göre sınıflamak amacıyla üretilen ağaç yukarıdan aşağı doğru incelendiği zaman toplam soru değerlendirme sayısı 38'den fazla olan ve tartışma ortamında cevap yazdığı soru sayısı 2'den fazla olan öğrencilerin %66.7 olasılıkla (n = 24) Başarılı sınıfında yer alacağı anlaşılmaktadır. Bunun aksine toplam soru değerlendirme sayısı 38'e eşit veya altında olup tartışma ortamında yazdığı cevap sayısı da 2'ye eşit veya daha az olan öğrencilerin %82.1 olasılıkla Başarısız sınıfında yer alacağı (n = 28) anlaşılmaktadır.

Kural tabanlı bir algoritma olan CN2 algoritması ile elde edilen sonuçlar incelendiğinde (Şekil 3.) başarısız ve başarılı öğrencileri sınıflamak amacıyla kullanılabilecek kurallar üretildiği görülmektedir. Şekil 4'de verilen bir numaralı kurala göre ortama giriş sayısı 17'ye eşit veya daha az olan ve öğrendiği kavramlarla ilgili yazdığı yansıma sayısı 12'ye eşit veya daha az olan öğrencilerin dersten başarısız olduğu görülmektedir (n = 13). Başarılı öğrencileri sınıflamak amacıyla kullanılabilecek dört numaralı kurala göre ise tartışma ortamında soru değerlendirme sayısı 38'e eşit veya fazla olan, tartışma ortamındaki sorulara yazdığı cevap sayısı 2 veya daha fazla olan ve yazdığı yansımalarda kullandığı etiket sayısı 11'e eşit veya daha fazla olan öğrencilerin dersi geçtiği görülmektedir (n = 20).

Tartışma, Sonuç ve Öneriler: Araştırma sonuçları öğrencilerin çevrimiçi öğrenme ortamındaki etkileşimlerini yansıtan bir takım verilerin kullanılarak derste sergileyecekleri akademik performanslarının önemli ölçüde tahmin edilebileceğini göstermiştir. En iyi sınıflama performansına sahip olan Naïve Bayes algoritması çapraz geçerlilik sonucu öğrencilerin %75.4'ünün dönem sonu performanslarını (Başarısız, Orta ve Başarılı) doğru olarak sınıflamıştır. Kaldı (Başarısız) – Geçti (Orta ve Başarılı) şeklinde bakıldığında ise modelin derste başarısız olan 27 öğrenciden 22'sini (%81.5), dersi geçen 49 öğrenciden ise 45'ini (%91.8) doğru olarak sınıfladığı görülmektedir. Diğer taraftan Karar Ağacı ve CN2 algoritmalarının ise uzman olmayan kişiler tarafından bile kolaylıkla yorumlanabilecek çıktılar ürettiği görülmüştür.

Bu araştırma göstermektedir ki öğrencilerin çevrimiçi ortama giriş sayıları, ortamda kalma süreleri, çevrimiçi tartışmalara katılmaları (cevap yazarak ya da soru ve cevapları değerlendirmek suretiyle), öğrendikleri kavramlarla ilgili yansıma yazmaları ve yazdıkları yansımaları kavramları kullanarak etiketlemeleri derste sergileyecekleri performans üzerine önemli etkisi olan değişkenlerdir. İleriki çalışmalarda öğretmenler bu bilgileri öğrencilere bireysel ve otomatik dönüt vermek amacıyla kullanabilir. Öğretim tasarımcıları çevrimiçi öğrenme ortamlarını bu bilgiler doğrultusunda düzenleyebilir, öğrencilerin bu değişkenler açısından katılımlarını artıracak tasarımlar yapabilirler.