CLASSIFICATION OF STUDENTS' ACHIEVEMENT VIA MACHINE LEARNING BY USING SYSTEM LOGS IN LEARNING MANAGEMENT SYSTEM

Dr. Ilhan KOYUNCU
ORCID: 0000-0002-0009-5279.
Faculty of Education
Adiyaman University
Adiyaman, TURKIYE

Dr. Abdullah Faruk KILIC
ORCID: 0000-0003-3129-1763
Faculty of Education
Adiyaman University
Adiyaman, TURKIYE

Dr. Derya ORHAN GOKSUN
ORCID: 0000-0003-0194-0451
Faculty of Education
Adiyaman University
Adiyaman, TURKIYE

Received: 26/06/2021 **Accepted:** 15/10/2021

ABSTRACT

During emergency remote teaching (ERT) process, factors affecting the achievement of students have changed. The purposes of this study are to determine the variables that affect the classification of students according to their course achievements in ERT during the pandemic process and to examine the classification performance of machine learning techniques. For these purposes, the logs from the learning management system were used. In the study, analyzes were carried out with various machine learning techniques and their performances were compared. As a result of the study, it was observed that Fisher's Linear Discriminant Analysis was the best technique in classification according to F measure performance criteria. As another result, the most effective variable, in classifying students, is the average number of days logged into the system per month and week. It has been observed that total activity duration (min), total number of weeks and total number of page views during the semester are less influential factors. Accordingly, it could be suggested to check the monthly and weekly follow-up of the lectures instead of the total follow-ups per semester. In addition, students' interaction patterns can be monitored with course tracking systems.

Keywords: Emergency remote teaching, linear discriminant analysis, machine learning, measurement and assessment, pandemic process, COVID-19.

INTRODUCTION

Covid-19 pandemic process caused to troubles at higher education institutions from various aspects as in numerous fields of life owing to restrictions. Face-to-face education systems were laid over and emergency remote teaching (ERT) activities were conducted at Turkiye along with other all countries. At this unpredictable process, with the intent of being able to continue learning activities, not disrupting instruction programs, typical face-to-face education elements like contents, teaching approaches, strategies and methods were presented to students via online platforms. Whereas distance education is required to different instructional design processes from face-to-face teaching. Because of instructor could not reach these design opportunities in current conditions, conducted activities were named as ERT (Hodges, Moore, Lockee, Trust and Bond, 2020).

Although National Ministry of Education in Turkiye have decided different strategies primary, secondary and high school stages occasionally, all universities have chosen ERT for theoretical courses even some of them made it for applied courses as well. Universities planned these processes with their own means and opportunities. Some of ways of ERT at universities are learning management systems (LMSs), video conferencing tools, Google tools, Web 2.0 tools. Even though each ERT ways of universities differ from others, at essential points of them have similarities such as accessing course materials, course presentations, participation to synchronous courses, online exams. These similarities give an opportunity to comparisons of activities and common assessments. There are several studies on literature which students and instructors' views on ERT activities, suggestions on the frame of course and field, achievement comparisons (Aboyinga and Nyaaba, 2020; Piilikangas and Lindfors, 2021; Yazawa; 2021). These studies are result-oriented and depend on self-reported data by a majority. In spite of several scientific findings are able to ascertained on the basis of self-reported data, these data are limited for predictions, creating generalizable findings (Brutus, Aguinis and Wassmer, 2013; Lauritsen, 1999).

Data about student-system interactions are recorded on LMSs and Web 2.0 tools. These data can be exemplified as participation frequency to discussions, the days which students logged into course page, how much time s/he passed time etc. The findings, revealed by worked of system logs which is one of the most important data sources for comprehensive analysis and synthesis and are able to give lots of information about process, are become popular and widespread on literature. These data, seen as independent and relevant of each other, is named as big data and give opportunities for revealing important findings via patterns which were gained from analysis.

The patterns, created by analyzing gathered complex data, are named as learning analytics. Learning analytics consist of three processes as preoperational process, post operation analysis and actions (Bahceci, 2015). Accordingly, learning analytics cover whole processes of gathering analyzing data, giving feedbacks to respective people, using these for development and updating. Briefly, learning analytics is an ever-developing field, which uses students online activities for rising learning and achievement (Saqr, Fors and Tedre, 2017). It was highlighted that students interactions can be monitored and made provisions against probability od failure (Karaoglan-Yilmaz, 2020). In this sense, learning analytics give opportunities for following student activities, examining instructional contexts and gaining important feedbacks.

When reviewing literature, both studies on describing learning analytics (Dietz-Uhler and Hurn, 2013; Elias, 2011) and predicting student achievement by using these analytics (Aydin and Ozkul, 2015; Fernandez-Delgado, Mucientes, Vazquez-Barreiros and Lama, 2014; Saqr, et al., 2017) were observed. The studies on educational data mining, published at between 2006 and 2016 were examined and it was reported by Tekin and Oztekin (2018) that most of them focused on academic achievement and conducted by literature review. Besides this, it was seen that there are studies, investigated effects of online learning environments on the basis of a course (Bulca and Demirhan, 2020). Moreover, it can be observed in an obvious way, learning analytics studies can be conducted on the basis of a course or on national level (Dietz-Uhler and Hurn, 2013).

When creating learning analytics, several data mining techniques were set on. These can be exemplified as text mining, social media analytics, machine learning techniques. Also, it was known that a transition process to semantic systems in education and pandemic conditions were speeded up (Devedzic, 2004). In general, it can be said that operations and progressions are realized via patterns, which specified by instructors, individualized learning environments at these days and future. In other words, instructors assign to students preconditions and context modules in a nonlinear progression. For example, the students, finished module A, can continue with modules B or C, the students, finished modules A and B, can access to module D or E, module F can be accessed without finishing module D. These descriptions and specifications are identified by instructors. But this access is identified via general algorithm specifications, student's content achievement and interactions in semantic learning environments (Ohler, 2008). At this point, algorithms can be created in the frame of patterns, gathered by machine learning techniques, and system achievement can be raised. These algorithms are set on patterns, gathered by analyzing big data consist of human-computer interactions, for developing systems, can think nearly human intelligence (El Naqa and Murphy, 2015). By this way, guardiancies are provided based on interaction data and content achievement, individualized effective learning environments are designed (Kivinen, Warmuth and Hassibi, 2006).

Machine learning is a computer science field, aim to ensure computer "learning" without direct programming (Samuel, 1959). It has developed from 1950s, based on artificial intelligence and deep learning initiative, till today and especially is formed within the frame of implemented targets particularly prediction and optimization (Bi, Goodman, Kaminsky and Leesler, 2019; van Ginneken, 2017). Computer learning in machine learning is developed via interactions and experiences within the context of a task (Mitchell, 1997:15). These interactions and experiences are based on data in implementation. Above mentioned data are named as big data and can be gathered from learning tools such as LMSs, video context, screens, social media platforms, interactive videos. Patterns are tried to create by analyzing big data via machine learning techniques. For this reason, there is no certain differentiations between machine learning and algorithmic statistic techniques such as least absolute shrinkage and selection operator (LASSO), stepwise regression (Bi, et al., 2019). It can be said in conclusion that machine learning is successful algorithms and patterns by analyzing big data, gathered from logs, via various statistical techniques.

The success of machine learning technique is important for creating patterns. The machine learning techniques can be exemplified as logistic regression, decision trees and Naive Bayes classifiers etc. in a general point of view (Mackenzie, 2015). These techniques have advantages and disadvantages from the point of findings. Besides each has different preconditions, they create predictions by using different formulizations. For these reasons, experimenting these techniques on the same data and comparing them by this way are important for reaching valid and reliable findings, would be used when designing instructional systems and context. The current study has one another originality about comparing different machine learning techniques on the data, gathered from students, had no alternatives apart from ERT.

Importance of the Study

In literature, there are studies implementations on distance education and ERT by using analytic sources besides face-to-face teaching (Karaoglan-Yilmaz, 2020; Saqr et al., 2017). Data of this study consist of students-system interaction logs, who are at university and take courses via ERT. From the point of this view, it can be said that analyzed data have a holistic characteristic contextually. On the other hand, this study has an originality in terms of that the most appropriate technique was chosen by conducting frequently used machine learning techniques along and thuswise incorrect and subjective findings could be eliminated. It was seen that the other studies, used these techniques, had an approach in a similar way (e.g., Kotsiantis, Pierrakeas and Pintelas, 2003; Osmanbegovic and Suljic, 2012; Romero, Espejo, Zafra, Romero and Ventura,2013). Moreover, within the context of measurement and evaluation course, students' the time (min) passed into LMS and number of logging to LMS were investigated on the basis of semester, mount, week and day. It was thought that investigation of these variables would be contribute to examining student achievement, controlling and updating compulsory ERT activities.

Purpose of the Study

According to previous explanations, the aim of this study was examining which variables could affect students' achievement during teaching activities by ERT. With this aim, the logs, which were gained from LMSs, used for measurement and evaluation course presented by ERT, were analyzed. The instructor shared with students 14 course content presentation videos, 606 minutes in total, at 14 week-period time at whole semester. Each of instructional contents' average duration was approximately 43 minutes, they were presented to students weekly. At the semester, two synchronous online courses that students must attend them and one synchronous online course for reviewing theoretical content and solving sample problems were performed. Besides these, the electronic document of relevant course content in pdf format was sent to students at each week. At the same time, each week, discussions page was created and students were encouraged to participate them. To achieve this aim, machine learning techniques were used and tried to answer the research questions listed below:

According to students' average number of page views (daily, weekly, monthly, and in total) in the learning management system, average number of days (weekly and monthly) and total number of weeks logged into the system, and total time (min) spent in the system;

1. Which machine learning technique has the highest performance in classifying students as "passed" and "failed"?
2. What are the classification results of the machine learning technique with the highest classification performance?
3. What is the importance level of all these variables in classification?

METHOD

In this study, students' average number of page views (daily, weekly, monthly, and in total) in the learning management system, average number of days (weekly and monthly) and total number of weeks logged into the system, total time (min) spent in the system were described, and hence it is a descriptive survey research. In such studies, it is aimed to describe the existing characteristics of students (Karasar, 2014). In addition, it is also a correlational research since it examines the relationship between the success-failure of students and their interactions in the learning management system. In correlational studies, it is aimed to explore the relationships between the variables (Buyukozturk, Kilic-Cakmak, Akgun, Karadeniz and Demirel, 2013).

Participants

The participants of this study are 284 students studying at the education faculty of a state university in Turkiye. The participants were selected according to purposeful sampling method. This sampling method allows in-depth examination of the cases that provide satisfying information about research problem (Buyukozturk et al., 2013). The distribution of the participants according to their departments and gender is given in Table 1.

Table 1. Number of students by department and gender

Department/Gender	Female		Male		Total	
	f	%	f	%	f	%
Guidance and Psychological Counseling	56	62,2	34	37,8	90	31,7
Social Sciences	42	79,2	11	20,8	53	18,7
Science Teaching	17	85,0	3	15,0	20	7,0
Early Childhood	33	80,5	8	19,5	41	14,4
Religion and Culture	54	67,5	26	32,5	80	28,2
Total	202	71,1	82	28,9	284	100,0

As seen in Table 1, 202 (71.1%) of the participants were female and 82 (28.9%) were male. 90 (31.7%) of the participants attending to Guidance and Psychological Counseling Program, 53 (18.7%) to Social Sciences Education, 20 (7.0%) to Science Education, 41 (14.4%) to Early Childhood Education, and 80 (28.2%) to Religion and Culture Education programs.

Data Collection and Analysis

Within the scope of the study, system log reports of the measurement and evaluation course in the faculty of education given in the form of ERT throughout the semester were examined. This course is given at all departments of education faculties and related departments of other faculties having pedagogical formation education. In order to obtain the data to be used in the present study, students' average number of page views (daily, weekly, monthly, and in total) in the learning management system, average number of days (weekly and monthly) and total number of weeks logged into the system, total time (min) spent in the system were evaluated. Descriptive statistics regarding these variables are presented in Table 4. Accordingly, the average number of page views was calculated as follows:

$$average\ page\ view_{daily} = \frac{total\ page\ views}{total\ number\ of\ days\ logged\ into\ the\ system} \ldots 1$$

$$average\ page\ view_{weekly} = \frac{total\ of\ average\ number\ of\ views\ per\ week}{total\ number\ of\ weeks\ logged\ into\ the\ system} \ldots 2$$

$$average\ page\ view_{monthly} = \frac{total\ of\ average\ number\ of\ views\ per\ month}{total\ number\ of\ months\ logged\ into\ the\ system} \ldots 3$$

The average number of days logged into the system was calculated as follows:

The average number of days logged into the system was calculated as follows:

$$average\ number\ of\ logins_{weekly} = \frac{total\ number\ of\ days\ logged\ into\ the\ system}{total\ number\ of\ weeks\ logged\ into\ the\ system} \ldots 4$$

$$average\ number\ of\ logins_{monthly} = \frac{total\ number\ of\ days\ logged\ into\ the\ system}{total\ number\ of\ months\ logged\ into\ the\ system} \ldots 5$$

In this study, machine learning techniques were used in order to examine how effective fail-pass decisions of the students are predicted by using the students' log reports (interaction activities) in a learning management system. In order to prevent erroneous and biased results, many techniques are used together and the most appropriate one is chosen in machine learning applications (e.g. Kotsiantis, et al., 2003; Osmanbegovic and Suljic, 2012; Romero, et al., 2013). Accordingly, the performance of a wide variety of well-known techniques is often tested and compared. Therefore, in this study, out of the 34 analysis techniques, it was aimed to select the technique with the highest classification performance. In Table 2, it is given that these techniques are Bayes (Naïve Bayes Simple, Naïve Bayes, BayesNet, and Bayesian Logistic Regression), discriminant analysis (FLDA, QDA, and LDA), logistic regression (Kernel, Logistic, and Simple), neural networks (RBF Network, RBF Classifier, MLP Classifier, WiSARD, and Multilayer Perceptron), decision trees (J48, J48 Consolidated, NB Tree, Random Tree, Decision Stump, Random Forest, Extra Tree, FT, REP Tree, BF Tree, LMT, and Simple CART) and instance-based [nearest neighborhood] (KStar, IBk, and RseslibKnn), and rule-based (Decision Table, PART, DTNB, and OneR). Detailed information on the working principles of these techniques is included in the related packages of WEKA (Hall et al., 2009) software.

Table 2. Machine learning techniques used

Functions	Bayes	Trees	Lazy	Rules
FLDA	Naïve Bayes Simple	J48	KStar	Decision Table
QDA	Naïve Bayes	J48 Consolidated	IBk	PART
RBF Network	BayesNet	NB Tree	RseslibKnn	DTNB
Kernel Logistic Regression	Bayesian Logistic Regression	Random Tree		OneR
Logistic		Decision Stump		
LDA		Random Forest		
RBF Classifier		Extra Tree		
MLP Classifier		FT		
WiSARD		REP Tree		
Multilayer Perceptron		BF Tree		
Simple Logistic		LMT		
		Simple CART		

Note: FLDA: Fisher's Linear Discriminant Analysis, QDA: Quadratic Discriminant Analysis, RBF: Radial Basis Functions, LDA: Linear Discriminant Analysis, MLP: Multilayer Perceptron, NB: Naïve Bayes, FT: Functional Trees, RP: Reduced-Error Pruning, BF: Best-First, LMT: Logistic Model Trees, CART: Classification and Regression Trees, PART: Partial Decision Trees, DTNB: Naive Bayes and Decision Tables, KNN: K-Nearest Neighborhood

As the dependent variable of the study, the students' end-of-term grade point averages from the assessment and evaluation course were used. These grades were calculated by summing 20% of the first and second midterm exam grades and 60% of the final exam grades. In order to pass a course at the state university where this study was conducted, the end-of-year passing grade must be at least 60 points. Therefore, the cut-off score was taken as 60, and those whose end-of-term average was 60 and above were coded as "passed", and those below 60 were coded as "failed". In order to facilitate the analysis and increase the optimization of some techniques, all numerical variables were standardized as 0 in mean and 1 in variance before analysis. The whole steps that followed in the current study are given in Figure 1.

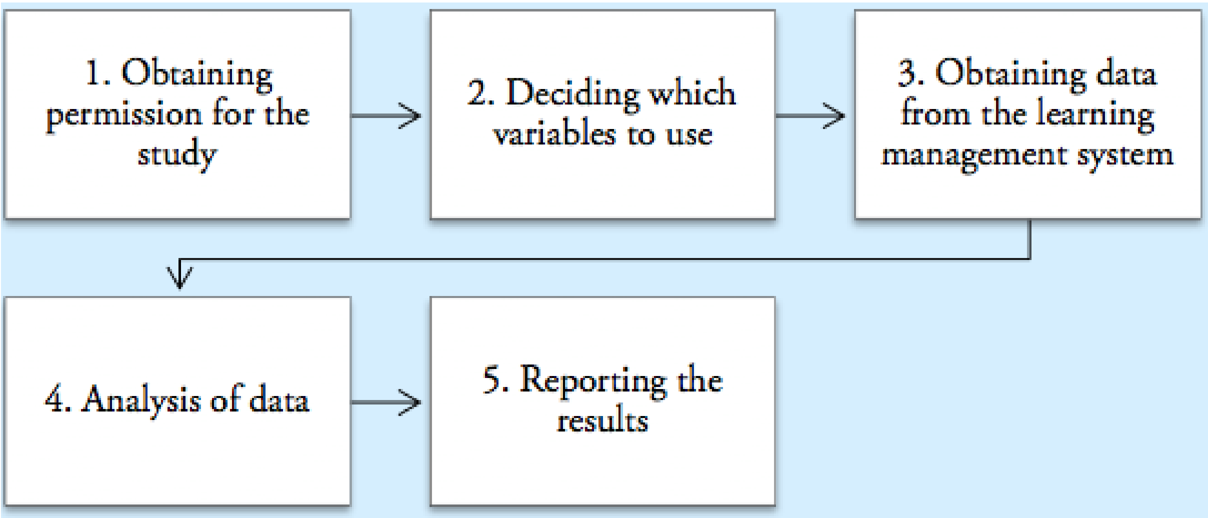


Figure 1. The Workflow of the Study

In evaluating the performance of machine learning techniques, threshold metrics (Accuracy, error rate, sensitivity, specificity, precision, recall, *F* measure, Kappa fit) and ranking metrics (Receiver operating characteristic [ROC] curve and precision-recall curve [PRC]) are mostly used. In determining which of these metrics to use, it is checked whether the number of observations in the dependent variable is balanced or not. In this study, there is an imbalanced dependent (class) variable ($n_{passed} = 215$, $n_{failed} = 69$). *F* measure is used as a model evaluation criterion in imbalanced distributions (Han, Pei and Kamber, 2011). In addition, since precision and recall values have equal importance for measurement and assessment course examined in the present study, F measure which was evaluated from equal weighting of these two measures will be one of the most appropriate evaluation metrics. Therefore, in this study, *F* values of the models were used as model evaluation criteria, and they were compared statistically with the dependent sample *t* test. In addition, the area under the ROC curve and accuracy values were also reported.

The techniques used to test the performance of models in machine learning are holdout, cross-validation and bootstrap methods. The holdout method, which is traditionally used in machine learning, is based on dividing the data into two parts, one part as train (usually 2/3 of data) and one part as test data (usually 1/3 of data) for analysis. Bootstrap method allows to obtain large samples by resampling from the existing data set. On the other hand, cross-validation allows the classification performance average to be obtained as a result of dividing the data into 10 equal parts and using one of them as test and the others as learning data in each iteration (Han, et al., 2011). In this study, all analyzes were performed with 10-fold cross-validation technique by using WEKA (Hall et al., 2009) software.

FINDINGS

In this section, findings of the study are presented under separate sub-headings for each research problem.

Selection of the Technique with the Highest Classification Performance

In Table 3, *F* measure, area under ROC curve, and percent correct values of machine learning techniques are given together with their standard deviation values. In addition, graphical representation of performance comparison of techniques is given in Figure 2.

Table 3. Classification performances of machine learning techniques

Techniques	F	Sx	ROC	Sx	PC	Sx
Fisher's Linear Discriminant Analysis (FLDA)	0.50	0.13	0.77	0.09	69.65	8.33
Naïve Bayes Simple (NBS)	0.49	0.14	0.74	0.10	70.32	9.16
Naïve Bayes (NB)	0.48	0.14	0.74	0.10	69.86	9.24
BayesNet (BayesN)	0.47	0.17	0.70	0.12	76.99	7.21
Quadratic Discriminant Analysis (QDA)	0.46	0.13	0.74	0.11	65.75	9.56
J48	0.46	0.19	0.66	0.13	77.06	7.47
J48Consolidated (J48_C)	0.46	0.14	0.67	0.13	65.75	10.00
Radial Basis Functions Network (RBFN)	0.42	0.19	0.73	0.11	78.12	5.99
KStar	0.41	0.15	0.66	0.11	72.45	7.28
Decision Table (DT)	0.41	0.19	0.66	0.12	77.83	6.58
Partial Decision Trees (PART)	0.41	0.21	0.68	0.11	75.91	7.03
Naive Bayes and Decision Tables (DTNB)	0.40	0.19	0.68	0.12	77.21	6.46
Naïve Bayes Tree (NBT)	0.40	0.20	0.64	0.12	76.50	7.01
IBk	0.39	0.15	0.60	0.10	70.32	7.58
Random Tree (RT)	0.39	0.16	0.60	0.10	70.70	7.80
Decision Stump (DS)	0.38	0.18	0.60	0.09	75.00	5.06
Bayesian Logistic Regression (BLR)	0.37	0.19	0.61	0.09	77.25	5.46
Kernel Logistic Regression (KLR)	0.37	0.19	0.76	0.10	76.97	5.60
Logistic	0.37	0.19	0.76	0.10	77.01	5.58
Linear Discriminant Analysis (LDA)	0.36	0.19	0.77	0.09	77.36	5.50
Random Forest (RF)	0.36	0.18	0.74	0.11	75.06	6.19
Radial Basis Functions Classifier (RBFC)	0.35	0.16	0.76	0.10	76.75	4.66
Multilayer Perceptron Classifier (MPC)	0.34	0.21	0.75	0.11	74.83	6.79
WiSARD	0.34	0.16	0.57	0.11	70.32	7.44
Extra Tree (ET)	0.34	0.17	0.57	0.11	67.71	8.00
Functional Trees (FT)	0.34	0.17	0.67	0.11	73.95	6.26
Reduced-Error Pruning Tree (REPt)	0.34	0.19	0.60	0.13	75.48	6.14
Multilayer Perceptron (MP)	0.33	0.20	0.74	0.11	74.61	6.64
Simple Logistic (SL)	0.33	0.18	0.75	0.10	76.52	5.05
Best-First Tree (BFt)	0.33	0.17	0.57	0.11	73.97	6.25
Logistic Model Trees (LMT)	0.33	0.19	0.75	0.10	76.28	5.37
Simple Classification and Regression Trees (CART)	0.31	0.17	0.55	0.09	74.57	4.50
Rseslib K-Nearest Neighborhood (KNN)	0.28	0.16	0.68	0.11	74.36	5.01
OneR	0.27	0.17	0.56	0.08	73.28	5.73

Note: ROC: Receiver Operating Characteristic, Sx: Standard deviation, PC: Percent Correct

Table 3 and Figure 2 show that the highest *F* measure values were obtained by Fisher's Linear Discriminant Analysis (FLDA). In order to examine whether this value differs statistically significantly from the *F* values of other techniques, a paired samples *t* test was performed. According to the *t* test results, FLDA technique has a significantly higher *F* value than all other techniques ($p<0.05$). Similarly, the area under the ROC curve of this technique (0.77) indicates also high classification performance. At the same time, this value is higher or equal to the values obtained from all other techniques. The percent correct classification obtained with the FLDA technique is 69.65%. However, although a correct classification close to 100% indicates that the method performs well, the interpretation of this value will lead to incorrect inferences, since the dependent variable used in this study is imbalanced. For this reason, the number of correctly classified and incorrectly classified students for each category and precision and recall values were given in Table 4.

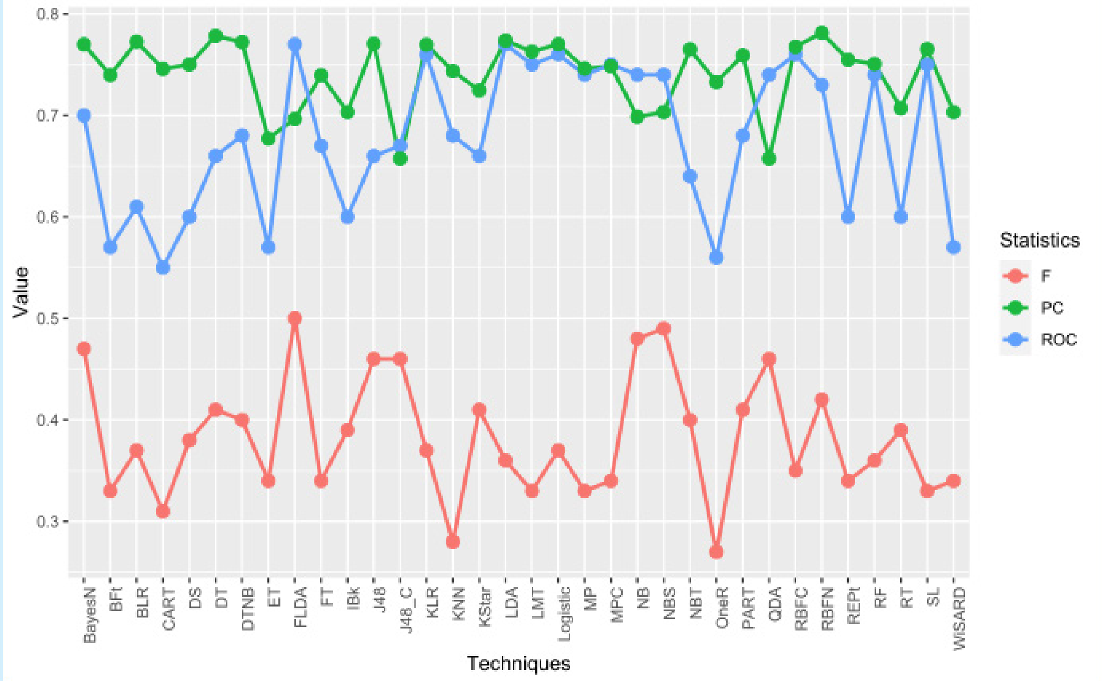


Figure 2. Performance Comparison of Machine Learning Techniques

Classification Results

In Table 4, the classification results obtained with the FLDA technique, which is the technique with the highest classification performance, are given.

Table 4. Classification results for FLDA technique

		Model Evaluation Criteria							Predicted			
		TP Rate	FP Rate	Precision	Recall	F Measure	ROC Area	PRC Area	Failed	Passed	Total	Percent
Actual	Failed	0.667	0.288	0.426	0.667	0.520	0.751	0.466	46	23		
	Passed	0.712	0.333	0.869	0.712	0.783	0.751	0.903	62	153		
Weighted Average		0.701	0.322	0.762	0.701	0.719	0.751	0.797			199	70.07

Note. TP: True Positive, FP: False Positive, ROC: Receiver Operating Characteristic, PRC: Precision-Recall Curve

In Table 4, it is seen that 199 out of 284 students were classified correctly according to the FLDA results. 23 students were classified as "passed" although they actually "failed" the course. On the other hand, 62 students were classified as "failed" although they actually "passed" the course. The ratio of students that were classified as "passed" although they actually "failed" is assumed to have equal importance with ratio of students that were classified as "failed" although they actually "passed" for the measurement and assessment course studied in the present study. Hence, *F* measure value were interpreted instead of precision and recall values. Accordingly, when *F* value (0.719) is interpreted together with the ROC value (0.751) and PRC value (0.797), it can be said that these values are close to 1 and the model has acceptable classification performance in general.

Variables in the Model

In Table 5, descriptive statistics and variable importance results regarding the independent variables used in the model are given.

Table 5. Variable importance results

Variables	Weight*	N	Min	Max	Mean	Sx
Average number of days logged into the system per month	0.624	284	2.60	21.20	8.25	3.24
Average number of days logged into the system per week	-0.514	284	1.13	6.24	2.55	0.83
average number of page views per month	-0.291	284	5.62	26.47	14.32	3.76
average number of page views per day	0.260	284	5.30	27.92	11.98	3.13
average number of page views per week	0.259	284	5.66	30.12	12.14	3.56
total time (min) spent in the system	-0.236	284	21.02	1424.32	297.85	194.30
total number of weeks logged into the system	-0.211	284	9.00	19.00	15.78	1.86
Total number of page views	-0.165	284	159.00	2066.00	476.04	213.19

* Threshold: 0.128, Sx: Standard deviation

In FLDA model, the weight values of the variables represent the coefficients in the model. The weight for each variable is interpreted as the power of that variable to separate "failed" and "passed" students. The sign of weight is used to obtain the discriminant function and assign students to the classes. If the values obtained for the students are equal or greater than threshold value, the students are assigned to class one, otherwise they are assigned to the other class. When the discrimination power of the variables is examined, it is seen that the most discriminating variable is average number of days logged into the system per month ($\bar{X}$ = 8.25 ± 3.24). The least discriminating variable is total number of page views ($\bar{X}$ = 476.04 ± 213.19). The other variables are listed from largest to smallest according to their discrimination power as follows: Average number of days logged into the system per week ($\bar{X}$ = 2.55 ± 0.83), average number of page views per month ($\bar{X}$ = 14.32 ± 3.76), average number of page views per day ($\bar{X}$ = 11.98 ± 3.13), average number of page views per week ($\bar{X}$ = 12.14 ± 3.56), total time (min) spent in the system ($\bar{X}$ = 297,85 ± 194,30), and total number of weeks logged into the system ($\bar{X}$ = 15.78 ± 1.86).

DISCUSSIONS AND CONCLUSION

In this study, the classification performance of machine learning techniques in classifying students as "passed" or "failed" were examined by using their log reports in the learning management system. As a result of the study, it was found that Fisher's Linear Discriminant Analysis (FLDA) have the highest classification performance according to F measure model performans criteria that used for imbalanced data. In addition, it has been observed that this technique makes more successful estimations according to area under ROC curve when compared to the other techniques. The FLDA technique is used in statistics, pattern recognition, and machine learning to find a linear combination of related variables to categorize events or objects into two or more categories (Li and Wang, 2014). This technique is also robust to violations of normality and homogeneity of variances assumptions when interactions between continuous independent variables do not affect the dependent variable (Knoke, 1982). Especially when the assumption of normality is provided in continuous data and the sample size is more than 50, the FLDA technique gives better results than many other used analyzes such as logistic regression (Pohar, Blas and Turk, 2004).

The discriminanting power of the variables used in the study is ranked from high to low as average number of logins (monthly and weekly), average number of page views (monthly, weekly and daily), total time (min) spent in the learning management system, the number of weeks logged into the system and total number of page views, respectively. Accordingly, it was observed that monthly, weekly and daily interactions in the learning management system are more powerful factors in discriminating successful and unsuccessful students when compared to average time spent per semester and average numbers of logins to the system in total. Koc (2017) found that students' participation in discussion and live class in the learning management system was positively related to their project and final scores. Since these activities are planned daily, weekly or monthly, it can be said that regular follow-up of the learning management system increases the success of the students. In another study in which monthly activities of students were observed throughout the semester, it was seen that 66% of students in the risk group who do not use the system regularly could not complete the course (Cohen, 2017). Emphasizing that following the learning activities of students in distance education gives teachers important information about the development of their students, Zhang and Almeroth (2010) developed the Moodog system for this purpose. This system not only provides teachers with important information, but also gives education researchers the opportunity to evaluate the usefulness of distance education systems.

Suggestions and Limitations

According to the study findings, since the most effective variable in the classification of students' achievement is the monthly average number of logins to the learning management system, it can be examined whether the students login to the system monthly or not. Similarly, weekly or even daily follow-ups should be done if possible. At this point, students who log into the system less than the average may be warned about this issue or encouraged to spend more time and log in. In addition, instead of examining the semesterly course activities of the students, it can be followed whether they attend the course on a monthly and weekly basis. Monitoring the changes in students' activities throughout the semester will contribute to identifying the students who are in the risk group in terms of dropping out, making the necessary interventions in the process, and monitoring the problem throughout the campus by the university authorities and initiating the necessary intervention programs (Cohen, 2017). In addition, students' interaction patterns can be monitored with course tracking systems such as Moodog (Zhang and Almeroth, 2010).

In this study, since Fisher's linear discriminant analysis (FLDA) was found to be the most effective technique in discriminating students according to the relevant variables, the FLDA technique may be preferred to the frequently used and known analyzes such as logistic regression, artificial neural networks, when working with continuous variables. According to Gao, et al. (2020) FLDA is a preferable technique for deep learning and machine learning systems. As known, deep learning is one of subfields of machine learning and blows up designing smart learning environments (Balyen and Peto, 2019). But deep learning methodology uses nonlinear transformations (Dargan, Kumar, Ayyagari and Kumar, 2020). From this point of view, it can be said that this study may be an inspiration for deep learning but it is not a pioneer, so these algorithmic techniques should be test for a deep learning environment.

It should be noted that this study is limited to the data of educational measurement and evaluation courses in faculty of education or faculties with pedagogical formation education. Instructional activities designed by the instructors in this course are another limitation of the study. Considering that the classified variable is performance in the course, the interaction of the students is directly related to the difficulty level of the course content and the content design of the instructor. For this reason, researchers can carry out similar studies by using teaching activities and different student analytics that they use for other courses. In addition, the logs kept in the learning management system, which is the source of the study, constitute the big data of the research. In this learning management system, no data can be kept on students' interactions with video content. Since the scope of big data will change in a course where student interactions can be recorded with video content presented from different learning management systems or Web 2.0 tools, different findings can be reached by comparison and relationship analysis.

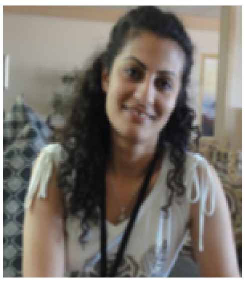
BIODATA and CONTACT ADDRESSES of AUTHORS

Ilhan KOYUNCU is a Doctor of Educational Measurement and Evaluation at the Education Faculty, Adiyaman University. Dr. Koyuncu gained his Ph.D. in Educational Measurement and Evaluation at February 2018. His academic interest areas are educational data mining, machine learning, multivariate statistics, scale development and adaptation, and exploratory and confirmatory factor analysis. He has over fifteen journal articles published in national and international indexes, five international book chapters, and papers submitted to national and international meetings.

Ilhan KOYUNCU
Department of Educational Sciences, Education Faculty
Address: Adiyaman University, 02000, Adiyaman, Turkiye.
Phone: +90 4162233800-1101
E-mail: ilhankync@Adiyaman.edu.tr

Abdullah Faruk KILIC is a Doctor of Educational Measurement and Evaluation at the Education Faculty, Adiyaman University. Dr. Kilic gained his Ph.D. in Educational Measurement and Evaluation at October 2019. His academic interest areas are dimensionality, exploratory and confirmatory factor analysis, item weighting, and reliability. He has over ten journal articles published in international indexes, four international book chapters, national and international articles, and papers submitted to international meetings.

Abdullah Faruk KILIC
Department of Educational Sciences, Education Faculty
Address: Adiyaman University, 02000, Adiyaman, Turkiye.
Phone: +90 4162233800-1119
E-mail: akilic@Adiyaman.edu.tr



Derya ORHAN GOKSUN got undergraduate degree at Firat University, 2010 and PhD degree at Anadolu University,2016, both were at Computer Education and Instructional Technologies Department, Faculty of Education. She got Associate Professor Doctor title at Computer Education and Instructional Technologies Scientific Field, 2020. Orhan Goksun, who worked for Adiyaman and Anadolu Universities from 2010 till today, is still working for Adiyaman University, Faculty of Education. Derya Orhan Goksun published her studies at national and international journals and presented at international many meetings, such as conferences, seminars, workshops. She studied as researcher at five scientific research projects supported by Adiyaman and Anadolu Universities and one project supported by TUBITAK (https://orcid.org/0000-0003-0194-0451) . 21st century skills, social networks, distance education, technology integration in education, learning theories, digital citizenship, technology use in special education, cyber behaviors, scientific research paradigms are some her research interests.

Derya ORHAN GOKSUN
Department of Instructional Technologies, Education Faculty
Address: Adiyaman University, 02000, Adiyaman, Turkiye.
Phone: +90 4162233800-1077
E-mail: dorhan@Adiyaman.edu.tr

REFERENCES

Aboyinga, J., & Nyaaba, M. (2020). Factors that ensure motivation in virtual learning among college of education students in Ghana: The emergency remote teaching (ERT) during Covid'19 pandemic. *European Journal of Research and Reflection in Educational Sciences Vol*, *8*(9), 1-9.

Aydin, S., & Ozkul, A. E. (2015). Veri madenciligi ve Anadolu Universitesi acikogretim sisteminde bir uygulama. *Egitim ve Ogretim Arastirmalari Dergisi*, *4*(2), 36–44.

Bahceci, F. (2015). Ogrenme yonetim sistemlerinde kullanilan ogrenme analitikleri araclarinin incelenmesi. *Turkish Journal of Educational Studies*, *2*(1), 41–58.

Balyen, L., & Peto, T. (2019). Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. *The Asia-Pacific Journal of Ophthalmology*, *8*(3), 264-272.

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). What is machine learning? A primer for the epidemiologist. *American Journal of Epidemiology*, *188*(12), 2222-2239.

Brutus, S., Aguinis, H., & Wassmer, U. (2013). Self-reported limitations and future directions in scholarly reports: Analysis and recommendations. *Journal of Management*, *39*(1), 48-75.

Bulca, Y., & Demirhan, G. (2020). Egitsel cevrimici sosyal ogrenme ortami EDMODO'nun fiziksel aktivite kavramini ogrenmede erisiye ve kaliciliga etkisi. *Egitim Teknolojisi Kuram ve Uygulama*, *10*(2), 577–589. doi:10.17943/etku.721876

Buyukozturk, S., Kilic-Cakmak, E., Akgun, O. E., Karadeniz, S., & Demirel, F. (2013). *Bilimsel arastirma yontemleri* (14th Edition.). Ankara: Pegem Akademi.

Cohen, A. (2017). Analysis of student activity in web-supported courses as a tool for predicting dropout. *Educational Technology Research and Development*, *65*(5), 1285-1304.

Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2020). A survey of deep learning and its applications: A new paradigm to machine learning. *Archives of Computational Methods in Engineering*, *27*(4), 1071-1092.

Devedzic, V. (2004). Education and the semantic web. *International Journal of Artificial Intelligence in Education*, *14*(2), 165-191.

Dietz-Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning*, *12*(1), 17–26.

Elias, T. (2011). Learning analytics : Definitions , processes and potential. *Learning*, 1–22.

El Naqa, I., & Murphy M.J. (2015) What is machine learning?. In: El Naqa I., Li R., & Murphy M. (Eds) Machine learning in radiation oncology. Springer, Cham. https://doi.org/10.1007/978-3-319-18305-3_1

Fernandez-Delgado, M., Mucientes, M., Vazquez-Barreiros, B., & Lama, M. (2014). Learning analytics for the prediction of the educational objectives achievement. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1–4). doi:10.1109/FIE.2014.7044402.

Gao, Z., Yuan, T., Zhou, X., Ma, C., Ma, K., & Hui, P. (2020). A deep learning method for improving the classification accuracy of SSMVEP-based BCI. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *67*(12), 3447-3451.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Peter, R., & Witten, I. H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations 11*(1), 10-18.

Han, J., Pei, J. & Kamber, M. (2011). *Data mining: Concepts and techniques*. Amsterdam, The Netherlands: Elsevier.

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause review*, *27*, 1-12.

Karaoglan-Yilmaz, F. G. (2020). Ogrenme analitigi geribildirimleri ile desteklenmis ters- yuz ogrenme ortaminin cesitli degiskenler acisindan modellenmes. *Bilgi ve Iletisim Teknolojileri Dergisi*, *2*(1), 1–16.

Karasar, N. (2014). *Bilimsel arastirma yontemi* (26th Edition). Ankara: Nobel Yayinlari.

Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2003). Preventing student dropout in distance learning using machine learning techniques. In *Knowledge-Based Intelligent Information and Engineering Systems* (pp. 267-274). Springer Berlin/Heidelberg.

Knoke, J. D. (1982). Discriminant analysis with discrete and continuous variables. *Biometrics*, 191-200.

Koc, M. (2017). Learning analytics of student participation and achievement in online distance education: A structural equation modeling. *Educational Sciences: Theory & Practice, 17*, 1893–1910. http://dx.doi.org/10.12738/estp.2017.6.0059

Kivinen, J., Warmuth, M. K., & Hassibi, B. (2006). The p-norm generalization of the LMS algorithm for adaptive filtering. *IEEE Transactions on Signal Processing*, *54*(5), 1782-1793.

Lauritsen, J. L. (1999). Limitations in the use of longitudinal self-report data: A comment. *Criminology*, *37*, 687.

Li, C., & Wang, B. (2014). *Fisher linear discriminant analysis*. CCIS Northeastern University. https://www.ccis.northeastern.edu/home/vip/teach/MLcourse/5_features_dimensions/lecture_notes/LDA/LDA.pdf

Mackenzie, A. (2015). The production of prediction: What does machine learning want? *European Journal of Cultural Studies, 18(4-5), 429–445.* doi:10.1177/1367549415577384

Mitchell, T.M. (1997). *Machine learning* (1st Edition) New York, NY: McGraw-Hill Education.

Ohler, J. (2008). The semantic web in education. *EDUCAUSE Quarterly*, *31*(4), 7-9.

Osmanbegovic, E., & Suljic, M. (2012). Data mining approach for predicting student performance. *Economic Review*, *10*(1), 3-12.

Piilikangas, A., & Lindfors, E. (2021). Stay online: Student teachers' views on online experiential learning in emergency remote teaching (ERT). *Techne Serien-Forskning i Slojdpedagogik och Slojdvetenskap*, *28*(2), 294-302.

Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: A simulation study. *Metodoloski Zvezki*, *1*(1), 143.

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.

Samuel A.L. (1959). Some studies in machine learning using the game of checkers. *IBM J Res Dev. 3*(3):210–229.

Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical Teacher*, *39*(7), 757–767. doi:10.1080/0142159X.2017.1309376

Tekin, A., & Oztekin, Z. (2018). Egitsel veri madenciligi ile ilgili 2006-2016 yillari arasinda yapilan calismalarin incelenmesi. *Egitim Teknolojisi Kuram ve Uygulama*, *8*(2), 67–89.

van Ginneken, B. (2017). Fifty years of computer analysis in chest imaging: Rule-based, machine learning, deep learning. *Radiological Physics and Technology*, *10*(1), 23-32. DOI: 10.1007/s12194-017-0394-5.

Yazawa, O. (2021) Comparison of Japanese university students' levels of self-determined motivation in face-to-face and ERT classrooms: NESTs and JTEs Effect on satisfaction of students' basic psychological needs. *Computer Assisted Language Learning Electronic Journal, 22*(2), 2021, 91-103

Zhang, H., & Almeroth, K. (2010). Moodog: Tracking student activity in online course management systems. *Journal of Interactive Learning Research, 21*(3), 407-429.