

Access this article online
Quick Response Code:

Website: www.jehp.net
DOI: 10.4103/jehp.jehp_394_18

Application of data mining techniques for predicting residents' performance on pre-board examinations: A case study

Leila Amirhajlou, Zohre Sohrabi, Mahmoud Reza Alebouyeh¹, Nader Tavakoli², Roghye Zare Haghghi³, Akram Hashemi⁴, Amir Asoodeh⁵

Abstract:

CONTEXT: Predicting residents' academic performance is critical for medical educational institutions to plan strategies for improving their achievement.

AIMS: This study aimed to predict the performance of residents on preboard examinations based on the results of in-training examinations (ITE) using various educational data mining (DM) techniques.

SETTINGS AND DESIGN: This research was a descriptive cross-sectional pilot study conducted at Iran University of Medical Sciences, Iran.

PARTICIPANTS AND METHODS: A sample of 841 residents in six specialties participating in the ITEs between 2004 and 2014 was selected through convenience sampling. Data were collected from the residency training database using a researcher-made checklist.

STATISTICAL ANALYSIS: The analysis of variance was performed to compare mean scores between specialties, and multiple-regression was conducted to examine the relationship between the independent variables (ITEs scores in postgraduate 1st year [PGY1] to PG 3rd year [PGY3], sex, and type of specialty training) and the dependent variable (scores of postgraduate 4th year called preboard). Next, three DM algorithms, including multi-layer perceptron artificial neural network (MLP-ANN), support vector machine, and linear regression were utilized to build the prediction models of preboard examination scores. The performance of models was analyzed based on the root mean square error (RMSE) and mean absolute error (MAE). In the final step, the MLP-ANN was employed to find the association rules. Data analysis was performed in SPSS 22 and RapidMiner 7.1.001.

RESULTS: The ITE scores on the PGY-2 and PGY-3 and the type of specialty training were the predictors of scores on the preboard examination ($F^2 = 0.129$, $P < 0.01$). The algorithm with the overall best results in terms of measuring error values was MLP-ANN with the condition of ten-fold cross-validation (RMSE = 0.325, MAE = 0.212). Finally, MLP-ANN was utilized to find the efficient rules.

CONCLUSIONS: According to the results of the study, MLP-ANN was recognized to be useful in the evaluation of student performance on the ITEs. It is suggested that medical, educational databases be enhanced to benefit from the potential of DM approach in the identification of residents at risk, allowing instructors to offer constructive advice in a timely manner.

Keywords:

Board certification examination, data mining, in-training examination, performance, preboard, prediction, resident

Departments of Medical Education, ³Deputy of Specialty and Subspecialty Education and ⁴Department of Medical Ethics, Iran University of Medical Sciences, Tehran, Iran, ¹Department of Anesthesiology and Pain Medicine, Iran University of Medical Sciences, Tehran, Iran, ²Department of Emergency Medicine, Iran University of Medical Sciences, Tehran, Iran ⁵Health Laboratories Administration, Birjand University of Medical Sciences, Birjand, Iran

Address for correspondence:

Ms. Leila Amirhajlou,
Department of Medical Education, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.
E-mail: l.amirhajlou@gmail.com

Received: 18-11-2018
Accepted: 01-02-2019

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Amirhajlou L, Sohrabi Z, Alebouyeh MR, Tavakoli N, Haghghi RZ, Hashemi A, *et al.* Application of data mining techniques for predicting residents' performance on pre-board examinations: A case study. *J Edu Health Promot* 2019;8:108.

Introduction

The field of educational data mining (EDM) has emerged as a result of the growing availability of educational data as well as the need to analyze large amounts of data generated from the educational ecosystem.^[1] EDM is a multidisciplinary field focusing on the development, study, and application of computerized methods to detect patterns in large collections of educational data which may be impossible to analyze.^[2]

The effort of educational institutions in developing countries has been focused on the generation of only facts and figures. However, it was reported that these simple facts and figures provide no assistance for educational institutions in improving educational settings. Today, more sophisticated and emerging trends are required to enable educational institutions to turn their data into valuable information.^[3] In the domain of medical education, data frequently collected from students' interactions, course information, and other academic sources (e.g., administration and curricula) is of such size and type that require special techniques to discover new knowledge.^[4]

Students of medical sciences are the future workers in the healthcare system, and thus, the quality of healthcare systems directly depends on the quality of medical science education.^[5] The residency training program is a highly critical and difficult course. Assessment of learning outcomes during residency is performed by internal evaluation and annual in-training examination (ITE) at the end of the residency course, and the performance of residents is assessed by the preboard qualifying and board certification examinations. ITEs (also known as promotion examinations) are held annually in the clinical disciplines of medical residency courses. These examinations are considered one of the most important summative examinations during this period. Residents can be promoted to the next PGY level based on these evaluations.^[6] Predicting residents' academic performance is critical for educational institutions because strategic programs can be planned to improve or maintain residents' performance during their period of education.^[7]

A commonly applied prediction method in medical education disciplines is the linear regression model, mainly due to its ease of construction and data interpretation.^[8-16] However, the interrelationship between variables and factors for predicting performance are complicated and nonlinear, and traditional linear models are simply inadequate for dealing with the modeling of data containing nonlinear characteristics. As a result, data analysis techniques require sophisticated algorithms to predict student performance. As a result,

traditional methods may not be directly applicable to these types of data and problems.^[17]

One line of research which is promising in this area is the application of DM techniques which can turn large datasets into useful information and knowledge. In general, DM allows us to conduct a predictive analysis which cannot be easily associated with a particular theory used for identifying variables regarding the study context.^[18] An educational DM can be implemented with numerous techniques. The use of these methods can lead to the discovery of several types of knowledge, such as association rules (ARs) as well as classification, clustering, and pruning the data. Several algorithms involved in the classification task are utilized to predict students' performance.^[19] Researchers in several studies have compared different DM techniques, including artificial neural network (ANN), Support vector machine (SVM), genetic algorithm (GA), and AR mining to predict student's final grades.^[20-23]

The present study aimed to apply DM techniques to predict the academic performance of residents based on their historical data in the residency training program. In this regard, previous studies employed the linear regression model despite being impractical in this area. In the present research, a more practical method was introduced and applied based on the prediction functions of DM algorithms. The multi-layer perceptron (MLP), SVM algorithms, and multiple regression were used on residents' data to predict their performance, and the most accurate model for future predictions of resident performance was proposed.

Participants and Methods

A retrospective cohort study was conducted on 841 residents entering the residency programs of ophthalmology, internal medicine, cardiology, general surgery, otolaryngology, and neurology at the Iran University of Medical Sciences, Tehran, Iran, between 2004 and 2014.

The present research was a descriptive cross-sectional pilot study in which the data belonging to residents and their outcomes were retrieved from the computerized database of the Department of Residency Training at the Iran University of Medical Sciences. The use of this de-identified database was approved by the Institutional Review Board of Residency Training, and thus informed consent was not required. Six out of 27 specialties, including internal medicine, ophthalmology, general surgery, neurology, cardiology, and otolaryngology were selected using convenience sampling. The research population comprises 1317 residents who had entered the six mentioned residency programs in 2004 and

graduated by 2014. Residents with a large proportion of missing data were excluded from the study, and for 67 residents (5%) with one missing value, the scores were replaced by the mean value. Thus, the final analysis was performed on the data of 841 residents.

Data collection instrument

The data collection instrument was a researcher-made checklist consisting of two parts; the first part examined the demographic data of the residents, and the second part asked for the residents' scores. Based on findings from the literature review and expert suggestions, the composite score of ITEs in PGY-1 to PGY-3 of 300 points, formative clinical assessment scores of 150 points, formative written examinations of 90 points, annual written examinations of 60 points, sex, and type of specialty training were selected as the independent predictors of performance on the PGY-4, i.e., preboard examination, scores.

Data analysis

Descriptive analysis was performed to identify the general characteristics of data. A one-way analysis of variance (ANOVA) with *post hoc* comparisons was used to determine whether there was a statistically significant difference in ITE mean scores across specialties. Moreover, multiple regression analysis was employed to investigate the predictive power of independent variables for preboard scores. In the next step, three common supervised classifiers, including MLP-ANN, SVM, and linear regression, were used to construct models predicting residents' preboard scores based on variables selected as described above. These algorithms extract information and infer patterns from the data in two phases of training and testing. In the first phase, the algorithm takes the values of independent variables (input) and the dependent variable (output) to learn the relationship between them to construct a prediction model. Afterward, in the testing phase, the constructed models are validated to predict the categories (class labels) of new data. In this study, 10 cross-validation methods were applied to validate the models, in which the dataset was randomly divided into 10 parts, with nine parts used for training and one-tenth reserve for testing. This procedure was repeated 10 times, each time reserving a different tenth for testing. The predictive performance of the three models was then compared. The root mean square error (RMSE) and MAE were utilized to evaluate the accuracy of the models. Since the classification algorithm is applied to categorical data, the residents' composite scores on ITEs were classified in four categories of excellent (≥ 245), good, (215–245) medium (185–215), and poor (≤ 185). Finally, the patterns discovered in the training phase were converted into confident classification rules by the frequent pattern-growth algorithm. The minimum

confidence of 0.3 and minimum support of 0.2 were set to select the rules. SVM and ANN significant parameters used during training are presented in Table 1. All the parameters were obtained using the trial-and-error method. Furthermore, the GA was utilized to improve the parameters in the ANN.

Statistical software

Data analysis was performed in RapidMiner 7.1.001 statistical software (RapidMiner, Dortmund, North Rhine-Westphalia, Germany) and SPSS 22 (IBM, Armonk, NY, USA).

Results

From the 841 residents, 67, 250, 54, 123, 88, and 259 were studying in the fields of neurology, internal medicine, otolaryngology, general surgery, ophthalmology, and cardiology, respectively. In this research, mean scores for all residents increased in each year of training. The mean score for PGY-1 residents was 220 (standard deviation [SD] = 15.5), whereas it was 224 (SD = 17.02), 233 (SD = 16.8), and 244 (SD = 17.2) for PGY-2, PGY-3, and preboard examinations, respectively. Tables 2 and 3 present the descriptive statistics of the mean scores of residents in PGY1–4 divided by specialty and sex.

According to one-way ANOVA, a significant difference was observed among the six specialties in terms of the mean scores of PGY-1 and preboard examinations ($P < 0.000$) [Table 4].

In this study, a multiple regression model was used to measure the relationship of independent variables (specialty, sex, and ITEs scores in PGY-1, PGY-2, and PGY-3) and the dependent variable (score of the preboard examination). As shown in Table 5, only three of the eight predictive variables showed significant results based on the results of beta weights, including

Table 1: The artificial neural network and support vector machine parameters applied during training

Parameter	Value
ANN	
Number of layers	1
Number of neurons	3
Training cycle	1000
Iteration	800
Learning rate	0.7
Momentum	0.48
SVM	
Kernel type	dot
C	3.48
Convergence epsilon	8.22
Max iteration	100000

ANN=Artificial neural network, SVM=Support vector machine

Table 2: Descriptive statistics -the mean score on in-training examinations by field of study and gender

Field	Gender	PGY-1 score	PGY-2 score	PGY-3 score	Preboard	Total
Internal medicine	Female (145, 58%)	223.4	230.4	236.2	245.8	234
	Male (105, 42%)	221.9	226.4	235.2	246.9	232.6
		Significant=0.4	Significant=0.04	Significant=0.66	Significant=0.61	Significant=0.33
Cardiology	Female (128, 49%)	226.5	219.6	229.2	241.7	229.2
	Male (131, 51%)	226.3	218.1	232.2	243.4	230.03
		Significant=0.93	Significant=0.52	Significant=0.14	Significant=0.42	Significant=0.53
Neurology	Female (45, 63%)	221.4	238.6	245.4	254.2	239.9
	Male (22, 37%)	216.8	229.4	241.4	251.3	234.7
		Significant=0.138	Significant=0.003	Significant=0.32	Significant=0.34	Significant=0.028
Surgery	Female (35, 29%)	210.1	215.4	225.8	240.1	222.8
	Male (88, 71%)	207.9	218.6	227.8	243.5	224.4
		Significant=0.52	Significant=0.26	Significant=0.48	Significant=0.27	Significant=0.4
Ophthalmology	Female (38, 44%)	219.1	223.5	236.1	247.5	231.5
	Male (50, 56%)	213.8	222.6	234.5	242.2	228.3
		Significant=0.09	Significant=0.78	Significant=0.63	Significant=0.18	Significant=0.16
ENT	Female (25, 46%)	208.6	230.9	239.6	250.1	232.3
	Male (29, 54%)	211.1	226.2	238.3	243.2	229.6
		Significant=0.54	Significant=0.23	Significant=0.73	Significant=0.21	Significant=0.3
Total	Mean	220.07	224.02	233.80	244.9	230.07

PGY=Postgraduate year

Table 3: Descriptive statistics - average score on postgraduate year-1 to postgraduate year-4 by field of study

Field	PGY-1 score	PGY-2 score	PGY-3 score	Preboard score	Total
Internal medicine	222.79	228.76	235.86	246.2928	233.4309
Cardiology	226.43	218.83	230.72	242.6202	229.6514
Neurology	219.91	235.59	244.13	253.2746	238.2272
Surgery	208.55	217.74	227.27	242.5179	224.0230
Ophthalmology	216.10	223.05	235.17	244.5591	229.7244
ENT	210.00	228.40	238.88	246.3778	230.9171
Total	220.07	224.02	233.80	244.9927	230.7251

PGY=Postgraduate year

Table 4: Analysis of variance mean differences between specialties

ANOVA	PGY-1 score	PGY-2 score	PGY-3 score	Preboard score	Total
F	34.16	20.6	13.15	5.02	21.87
Significant	0.000	0.000	0.000	0.000	0.000

PGY=Postgraduate year

Table 5: Multiple linear regression of predictive variable on preboard results

Model	Unstandardized coefficients		Standardized coefficients	t	Significant
	B	SE			
Model 1					
Constant	143.975	11.880		12.119	0.000
Gender	0.418	1.134	0.012	0.368	0.713
Filed	-0.609	0.310	-0.069	-1.961	0.050
PG2 score	0.102	0.038	0.102	2.696	0.007
PG3 score	0.285	0.035	0.278	8.125	0.000
PG1 score	0.063	0.038	0.058	1.645	0.100

SE=Standard error

PGY-2 (B = 0.102, P = 0.007) and PGY-3 (B = 0.278, P = 0.000) scores and the type of specialty training (B = -0.069,

P = 0.050). On the other hand, no significant impact was observed on preboard scores for the variables of PGY-1 scores (0.100) and sex (0.713). The model summary revealed the R² = 0.129 (12.9%) of the variance explained by the predictors of the variables (F = 24.652; df = 5.836; P = 0.000 or P < 0.05).

The PGY-2 and PGY-3 scores and type of specialty training had a low correlation with preboard examination scores. Nevertheless, as a consequence of the nonlinearity of some variables, the total model failed to reach statistical significance to explain the variability in preboard examination scores. Therefore, this study also considered a comparative analysis of three machine learning algorithms, including MPL-ANN, SVM, and linear regression to predict residents' performance on the preboard examinations.

Table 6 represents the error measures of all algorithms, indicating that the MLP-neural network has a minimum error based on RMSE and MAE from among other algorithms. This result demonstrated that the MLP neural network model is able to correctly predict the performance

of residents on preboard exams (RMSE = 0.091, MAE = 0.061 condition: 10-fold cross-validation).

In the final step, the MLP neural network was applied to derive efficient rules from the data using the selected attributes, including the type of specialty training, sex, and the mean score on the ITE exams. Eventually, 49 rules were generated for minimum confidence (0.3) and minimum supports (0.2). Table 7 presents information about the top seven created rules. Two rules were selected based on expert opinion as follows:

1. 1.If the field of the resident is surgery and the average score is 185–215, then the sex is male, i.e., male surgery residents obtain lower scores than residents in other disciplines
2. 2. If the average score of the resident is >245 and the field of the resident is internal medicine, then the gender is female, i.e., female internal medicine residents obtain higher scores than residents in other disciplines.

Discussion

To the best of the author’s knowledge, this paper was the first study to assess the important associations between the scores of residents on ITEs. According to the results, the MLP neural network has a higher accuracy in designing the residents’ performance prediction system compared to linear regression and SVM. In this regard, our findings are consistent with results reported by Ajiboye and Arshaa^[23] showing that the neural network provided superior prediction results. However, our results are not in line with those obtained by Almarabeh^[20] demonstrating that the Bayesian network classifier had the highest accuracy from among various classifiers. In addition, Strecht *et al.*^[22] concluded that the SVM algorithm had the best results from among the evaluated methods. Finally, Depren *et al.*^[21] reported

that logistic regression outperformed other algorithms in terms of performance classification.

According to the results of the present research, there was a predictable correlation between residents’ PGY-2, PGY-3, and preboard scores with the duration of residency, which is in congruence with prior studies. For instance, Grossman *et al.*^[24] found that performance on the ITE was accurately predicted and highly correlated with performance on the ABIMCE. Other studies in several specialties (e.g., internal medicine, family medicine, orthopedic surgery, obstetrics/gynecology, psychiatry, and ophthalmology) demonstrated a positive relationship between ITE performance and future performance on board certification examinations over time (8; 9; 10, 11, 12,13; 14; 15; 16). In our study, based on the results of AR mining, a female resident with the score of > 245 would belong to the discipline of internal medicine. On the other hand, a male resident studying in the field of surgery would obtain a score within the range of 215–245. In addition, multivariate analysis indicated that the residents’ discipline was the only demographic characteristic to reach statistical significance as the predictor variable for the score of preboard examination. Our results revealed that general surgery residents had significantly lower levels of mean scores compared to other residents.

However, using multiple regression analysis, no significant association was observed between residents’ sex and performance on preboard examinations, which is consistent with the results obtained by Brateanu *et al.*^[12] Nevertheless, Stohl^[25] introduced sex as a reliable predictor for the performance of residents. In the present study, multiple linear regression results indicated that the total results of ITEs predicted only 12.9% of the variance in the scores of preboard examinations. This low percentage of explained variance indicated that the predictor variables were not adequate for explaining our model. There are several possible explanations for this low coefficient. First, “variance explained” may not be the best metric for interpreting the predictive ability of test scores. As discussed earlier, the linear regression model achieved a lower precision in predicting the residents’ performance on preboard exams compared to ANN and SVM. Second, we

Table 6: Accuracy of predictions of the three models

Algorithms	RMSEA	MAE
MLP-ANN	0.091	0.061
SVM	0.208	0.16
Linear regression	0.115	0.219

MLP-ANN=Multi-layer perceptron artificial neural network, SVM=Support vector machine, MAE=Mean absolute error, RMSEA=The Root Mean Square Error of Approximation

Table 7: Results of association rules mining

Rules	Support	Confidence	Lift
If 185 <average score <215, field of resident=cardiology>⇒gender=female	0.092857	0.534247	1.081367
If average score >245, field of resident=cardiology⇒gender=male	0.04881	0.569444	1.128145
If 185<average score <215, field of resident=internal medicine⇒gender=female	0.07619	0.598131	1.210674
If field of resident=surgery, gender=male⇒185 <average score <215	0.063095	0.602273	1.157687
If average score >245, field of resident=internal medicine⇒gender=female	0.082143	0.610619	1.235953
If field of resident=cardiology and gender=female⇒185 <average score <215	0.092857	0.614173	1.180562
If 185<average score <215 and field of resident=surgery⇒gender=male	0.063095	0.679487	1.346154

employed the total score of ITEs for predicting residents' performance based on insufficient recorded data. In the present study, not all residents' scores on PGY-1-PGY-4 and board certification were available. In addition, no subscores of ITEs, including internal written or verbal exam, mini-clinical evaluation exercise, directly observed procedural skills, chart stimulated recall, objective structured clinical examination, or professionalism scores was applied in this study.

Conclusions

This study utilized techniques including ANN, SVM, and linear regression for predicting residents' performance on the preboard examination. It can be concluded that ANN has superior results in terms of less error compared to other methods. It is expected that EDM should offer benefits for residents and medical schools. We hope that our results will encourage other researchers to apply more sophisticated DM techniques to assess the possibility of additional predictors which were not included in the present study due to insufficient recorded data, such as predicting board certification results based on all component scores of ITEs. Therefore, it is recommended that information technology solutions be applied for the medical education centers to systematically collect all residents' data, including low- and high-stakes assessments. Furthermore, it is suggested that the existing systems be upgraded to utilize prediction algorithms for forecasting performance and making decisions about residents' progress during the residency training program. In addition, a major drawback of the present study was including only six residency programs. Consequently, future studies must evaluate students' performance from all residency programs.

Acknowledgment

The authors would like to acknowledge the contribution of the Educational Department of Residency Program at Iran University of Medical Sciences and Dr. Kamran Soltani Arabshahi for their assistance in conducting this research.

Financial support and sponsorship

Iran University of Medical Sciences.

Conflicts of interest

There are no conflicts of interest.

References

- Al-Razgan M, Al-Khalifa AS, Al-Khalifa HS. Educational Data Mining: A Systematic Review of the Published Literature 2006-2013. Lecture Notes in Electrical Engineering Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013); 2013. p. 711-9.
- Romero C, Ventura S. Educational data mining: A survey from 1995 to 2005. *Expert Syst Appl* 2007;33:135-46.
- Rahman SA. Applicability of Educational Data Mining in Afghanistan: Opportunities and Challenges. In: Proceedings of the 9th International Conference on Educational Data Mining, 634-35. Raleigh, NC, USA; Available from: http://www.educationaldatamining.org/EDM2016/proceedings/paper_1.pdf. [Last accessed on 2017 Mar 02].
- Vaitsis C, Nilsson G, Zary N. Visual analytics in healthcare education: Exploring novel ways to analyze and represent big data in undergraduate medical education. *PeerJ* 2014;2:e683.
- Bahadori M, Mousavi SM, Sadeghifar J, Haghi M. Reliability and performance of SEVQUAL survey in evaluating quality of medical education services. *Int J Hosp Res* 2013;2:39-44.
- Yaghmaei M, Heidarzadeh A, Jalali MM. Relationship between residents' success on the certifying examinations with in - Training exam and internal evaluation. *Res Med Educ* 2017;9:26-18.
- Rusli NM, Ibrahim Z, Janor RM. Predicting Students' Academic Achievement: Comparison Between Logistic Regression, Artificial Neural Network, and Neuro-Fuzzy. *International Symposium on Information Technology*; 2008.
- Bedno SA, Soltis MA, Mancuso JD, Burnett DG, Mallon TM. The in-service examination score as a predictor of success on the American board of preventive medicine certification examination. *Am J Prev Med* 2011;41:641-4.
- Kay C, Jackson JL, Frank M. The relationship between internal medicine residency graduate performance on the ABIM certifying examination, yearly in-service training examinations, and the USMLE step 1 examination. *Acad Med* 2015;90:100-4.
- Hauer KE, Vandergrift J, Hess B, Lipner RS, Holmboe ES, Hood S, *et al.* Correlations between ratings on the resident annual evaluation summary and the internal medicine milestones and association with ABIM certification examination scores among US internal medicine residents, 2013-2014. *JAMA* 2016;316:2253-62.
- O'Neill TR, Li Z, Peabody MR, Lybarger M, Royal K, Puffer JC, *et al.* The predictive validity of the ABFM's in-training examination. *Fam Med* 2015;47:349-56.
- Brateanu A, Yu C, Kattan MW, Olender J, Nielsen C. A nomogram to predict the probability of passing the American board of internal medicine examination. *Med Educ Online* 2012;17:18810.
- Caffery T, Fredette J, Musso MW, Jones GN. Predicting American board of emergency medicine qualifying examination passage using United States medical licensing examination step scores. *Ochsner J* 2018;18:204-8.
- Althouse LA, McGuinness GA. The in-training examination: An analysis of its predictive value on performance on the general pediatrics certification examination. *J Pediatr* 2008;153:425-8.
- Norcini JJ, Grosso LJ, Shea JA, Webster GD. The relationship between features of residency training and ABIM certifying examination performance. *J Gen Intern Med* 1987;2:330-6.
- Tian C, Gilbert DL. Association between performance on neurology in-training and certification examinations. *Neurology* 2013;81:1102.
- Hughes G, Dobbins C. The utilization of data analysis techniques in predicting student performance in massive open online courses (MOOCs). *Res Pract Technol Enhanc Learn* 2015;10:10.
- Yukselturk E, Ozekes S, Türel YK. Predicting dropout student: An application of data mining methods in an online education program. *Eur J Open Distance E Learn* 2014;17:118-33.
- Juneja S. Research survey of data mining techniques in educational system. *Int J Eng Comput Sci* 2016;5:19010-13. [doi: 10.18535/ijecs/v5i11.50].
- Almarabeh H. Analysis of students performance by using different data mining classifiers. *Int J Mod Educ Comput Sci* 2017;9:9-15.
- Depren SK, Aşkın ÖE, Öz E. Identifying the classification

- performances of educational data mining methods: A case study for TIMSS. *Theory Pract* 2017;17:1605-23. [doi: 10.12738/estp.2017.5.0634].
22. Strecht P, Cruz L, Soares C, Mendes-Moreira J, Abreu R. A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance. *The 8th International Conference on Educational Data Mining (EDM 2015)*; 2015. p. 392-5.
 23. Ajiboye AR, Arshaa RA. A novel approach to efficient exploration and mining of students' data. *J Theor Appl Inf Technol* 2015;79:176-84. Available from: <http://www.jatit.org/volumes/Vol79No1/18Vol79No1.pdf>. [Last accessed on 2015 Nov 10].
 24. Grossman RS, Fincher RM, Layne RD, Seelig CB, Berkowitz LR, Levine MA, *et al.* Validity of the in-training examination for predicting American board of internal medicine certifying examination scores. *J Gen Intern Med* 1992;7:63-7.
 25. Stohl HE, Hueppchen NA, Bienstock JL. Can medical school performance predict residency performance? Resident selection and predictors of successful performance in obstetrics and gynecology. *J Grad Med Educ* 2010;2:322-6.