Indian Journal of Science and Technology, *Vol 11(7), DOI: 10.17485/ijst/2018/v11i7/119062, February 2018*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Predicting Teachers' ICT Competence in a Philippine University Using J48 Algorithm

Las Johansen B. Caluza*

Information Technology and Computer Education Department, Leyte Normal University, 6500 Leyte, Philippines; lasjohansencaluza@lnu.edu.ph

Abstract

Objectives: The advent of information technology causes a significant impact on pedagogy and using ICT technologies in school-related activities. With this argument, determining the teacher's ICT competence is expected to yield a better or positive impact on the learners and the performance of the teachers itself, however, no model was developed in predicting teachers ICT competence. In line with this, it is empirical to study and develop a model that will help the university in predicting teacher's ICT competence. **Methods:** Data mining approach utilizing J48 algorithm was applied in this paper to create a model suitable for the actual teachers' characteristics in the University. Moreover, Cross-validation technique was used to validate the dataset to have an optimum and acceptable model and generating the Receiving Operating Characteristics Curve (ROC) Area under ROC Curve technique. **Findings:** Decision tree model and decision rule for classification were created. Additionally, there were 92.78% correctly classified with an AUC weighted mean of 92.4%. Also, the model has very high acceptability and accuracy in predicting Teacher's ICT Competence. However, it also revealed that many teachers still need more exposure in utilizing ICT technologies in pedagogy and any school-related activities. **Application/Improvements:** The result of this study can be a basis for developing software that will automatically categorize or classify the Teacher's ICT competence. For more improvement of this paper and the model, it is suggested to add additional parameters to have more factors involved in predicting Teacher's ICT competence.

Keywords: Blended Learning, Decision Tree, ICT Competency, ICT Domain, ICT Pedagogy, J48 Algorithm, Machine Learning

1. Introduction

Universities around the world are continuing its effort to be at par with their counterpart about educational qualifications of their Professors. Educational qualification of the teacher is a factor of the success of the students and the implementation of an academic program of an institution. Many research outputs support this claim, like teaching staff who earned advanced degrees has a positive impact on their performance as teachers and influence students in achieving their goals[1]. In a separate study also revealed that teachers' educational background has a positive effect on teachers' degree type on the student achievement[2]. In fact, leaders of higher education were challenged to position their institutions to satisfy the connectivity needs of prospective students and meet growing expectations and requirements for high-quality learning experiences and consequences[3]. Passed on the increasing indication that Internet data and communication technologies are transforming much of society, there is little cause to think that it will not be the defining transformative innovation for higher.

In recent years, many universities introduced a new learning environment called Blended Learning (BL). BL is the mixing and matching of Face-to-Face, E-Learning, and Self-paced learning[4]. Masie described BL as a mixture of e-learning and classroom teaching which the teacher has specific allotted time to each learning environment[5]. In this case, technical skills in handling blended learning; especially in using ICT tools (hardware and software) are very vital. A study about ICT in Educational Policies in the Asian Region noted three main challenges based on the recent

**Author for correspondence*

empirical research such as closing the digital divide, promoting the safe and responsible use of ICT, and measuring and evaluating ICT literacy skills[6]. Moreover, a study on ICT use in the Philippine public and private schools revealed some inadequacies like the absence of information on how to use ICT, lack of harmonization among public and private sector efforts, and insufficient teacher preparation[7]. As a result, it hampers the teaching and learning process of the students and teachers in the 21st Century schools. In the same research, the Philippines and other developing countries in Asia are into educational technology specifically in ICT hoping to have a significant change in their educational system through pedagogical benefits associated with ICT integration[8]. Moreover, to produce ICT professionals and ICT-literate workforce, the Philippine Congress contemplated laws in integrating ICT into the curriculum at all levels of education[7]. Much research was conducted to unveil the status of ICT competencies of Teachers in the Philippines. Some results revealed varied levels and needs of the Teachers. An example of this research output says ICT competency level of the respondents were knowledge deepening level where the lowest level of competency is in ICT tools and cooperation, while ICT educational policy is understanding level8. In the same research, ICT integration in curriculum and assessment reportedly to be as knowledge application, while pedagogical integration of ICT as complex problem-solving. Additionally, arguably, relationships between educational qualification and ICT competency were among the most studied variables in research. However, what remains hidden is a model that will determine the teachers in handling or to handle a blended learning environment. This paper tends to look into the educational qualifications, ICT competencies, and other personal characteristics of the Professors to develop a model using a decision tree.

2. Methodology

This study anchors on the Knowledge Discovery in Database (KDD) from the work of Fayyad, Piatetsky-Shapiro, and Smyth[9,10]. The KDD procedure is interactive and iterative (with many choices made by the user), involving numerous steps, summarized as:

2.1 Variables Used

Age: is measured in years as the length of time that a person has lived or existed.

Domain A (DA): Technology Operations and Concepts. This domain includes competencies related to technical operations and concept, and productivity of various ICT tools like computers and communication devices as well as an application available online or off-line10.

Highest Educational Attainment (HEA): is defined as the last school degree attended and or graduated. The categorization of HEA is doctoral degree graduate, with doctoral units, master's degree graduate, with master's units, bachelor's degree graduate.

Internet Access (IA): availability and accessibility of internet in school, home, internet station, and mobile data.

Length of Service (LOS): the number of years in teaching in the university.

Level of Computing (LC): refers to the digital competency level of the teacher and being measured through pedagogical indicators and domain A.

Pedagogical Indicator (PI): This domain includes competencies related to the use of technology (National ICT Competency Standards for Teachers – Slideshare, ND):

Seminars Attended (SA): these are seminars, symposiums and the like relative to information technologies attended by the respondents.

Sex: is identified as male (M) and female (F).

Technology Resources (TR): these are resources available and being used for instruction and development of instructional materials. TR is categorized as with Hardware Resources (HR) and with Software Resources (SR).

Training Attended (TA): these are training, workshops and the like relative to information technologies attended by the respondents.

2.2 Machine Learning Application

Supervised Machine Learning (SML) is one of the most popular applications of machine learning. SML is a machine learning task used to infer labeleddatasets[11]. An example of SML task is the pattern classification tasks. In this example, predictive modeling is the general concept of constructing a model that is capable of making predictions, in which such a model includes a machine learning algorithm that finds out specific properties from a training dataset to get those predictions[12]. In this paper, a type of a supervised machine learning called Decision Tree was utilized to classify the characteristics of the Professors to unravel and developed a decision tree model.

2.3 Decision Tree

Decision trees are standard supervised learning algorithms, easy to understand and easy to use[13]. Decision trees are trees that separate instances by sorting them based on feature values. Each node in a decision tree embodies a characteristic in an occurrence to be classified, and each branch exemplifies a value that the node can assume. Cases are classified starting at the beginning node and sorted based on their characteristic values[14]. Decision trees are presently one of the most common methods used for data modeling. They have the advantage of being theoretically pure and have been shown to perform well on a variety of glitches. Decision trees have numerous uses, such as, for instance, predicting a likely result, supporting the analysis of problems, and aiding in making decisions. When articulating and configuring decision trees, the results of real-world factors are analyzed and compiled, such that the specifics of the foregoing factors and related results are used to predict the results of future factors as shown in Figure 1[15].

3. The J48 Algorithm

Classification is the method of building a model of classes from a set of proceedings that contain the class labels. J48 is an extension of ID3 with features such as accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, and others[16].

4. Cross-Validation

Cross-validation is a model evaluation method where the entire data will not be utilized when training a learner. The dataset is split into two, such as train dataset for training and test dataset for testing the performance of the learned model. In this paper, the researcher utilized the k-fold cross validation using 10-fold cross-validation. K-fold cross-validation is one way to improve the holdout method. The data set is divided into k subsets, and the holdout method is repeated k times. Each time, one of the k subsets is used as the test set, and the other k-1 subsets are put together to form a training set. Then the average error across all k trials is computed. The advantage of this method is that it matters less how the data gets divided. Every data point gets to be on a test set exactly once and gets to be in a training set k-1 times. The variance of the resulting estimate is reduced ask is increased[17].

4.1 Model Accuracy Using ROC Curve and AUC Curve

Receiving Operating Characteristics Curve (ROC) Area Under ROC Curve technique is a universal biostatistical tool for describing the accuracy of a model regarding predicting a phenomenon[18]. It is a plot of the Sensitivity (TPR) against the Specificity (FPR) for the different possible cutpoints of an analytical test[19]. ROC curve proves several things[20,21]:

- It shows the tradeoff amongst True Positive Rate (TPR) and False Positive Rate (FPR).
- The nearer the curve follows the left-hand border and then the top edge of the ROC space, the more accurate the test.
- The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test.

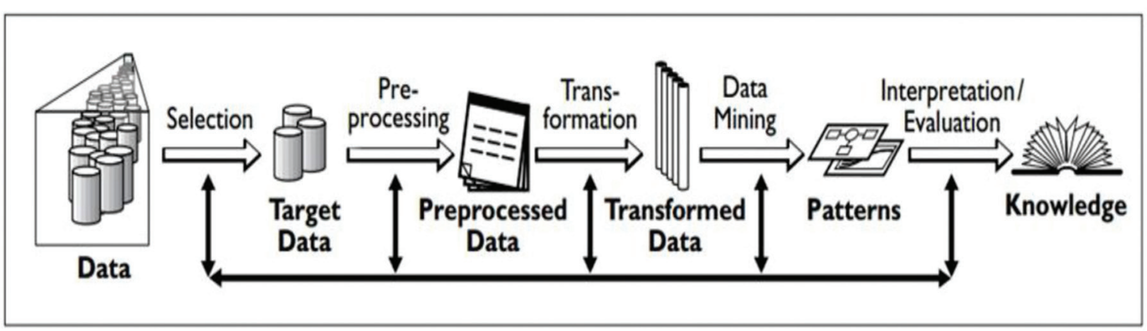


Figure 1. Overview of the steps constituting the KDD process.

5. Results and Discussion

5.1 The Model

Figure 2 is the graphical presentation of the pruned decision tree of Teacher's ICT Competence. Figure 3 showed Pedagogical Indicator (PI)as highly correlated and with the highest information gain and became the basis for its first split between PI and DA_S3 in predicting the Level of ICT Competency (LC) of the Teachers.

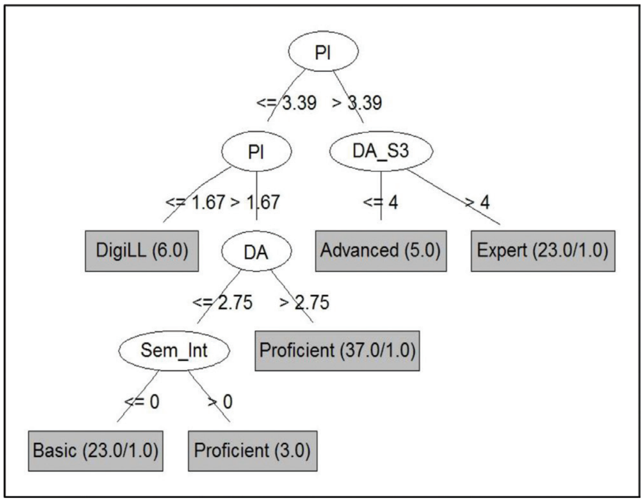


Figure 2. Teacher's ICT competence decision tree.

R_1: IF (PI <= 3.39) AND(PI <= 1.67) THEN
LC = "Digital Illiterate"
R_2: IF (PI <= 3.39) AND (PI > 1.67) AND (DA <= 2.75) AND (Sem_Int <= 0) THEN
LC = "Basic"
R_3: IF (PI <= 3.39) AND (PI > 1.67) AND (DA <= 2.75) AND (Sem_Int > 0) THEN
LC = "Proficient"
R4: IF (PI <= 3.39) AND (PI > 1.67) AND (DA > 2.75) THEN
LC = "Proficient"
R_5: IF (PI > 3.39) AND (DA_S3 <= 4) THEN
LC = "Advanced"
R_6: IF (PI > 3.39) AND (DA_S3 > 4) THEN
LC = "Expert"

Figure 3. Teacher's ICT competence decision rule.

The confusion matrix shown in Table 1 illustrates the correctly classified instances and the misclassifications of Teacher's ICT Competence. Moreover, the Confusion Matrix is interpreted as:

- The decision tree has classified 22 Experts objects as Experts and two as Advanced, leading to 2 misclassifications.
- The decision tree has classified 38 Proficient objects as Proficient, two as Advanced, leading to 2 misclassifications.
- The decision tree has classified 22 Basic objects as Basic, one as Proficient, and one as Digitally Illiterate, leading to 2 misclassifications.
- The decision tree has classified 3 Advance objects as Advanced and one as Proficient, leading to 1 misclassification.
- The decision tree has classified 5 Digitally Illiterate objects as Digitally Illiterate.

Table 1. Confusion matrix

Expert	Proficient	Basic	Advanced	DigiLL	<-- Classified as:
22	0	0	0	0	Expert
0	38	1	1	0	Proficient
0	0	22	0	0	Basic
2	2	0	3	0	Advanced
0	0	1	0	5	DigiLL

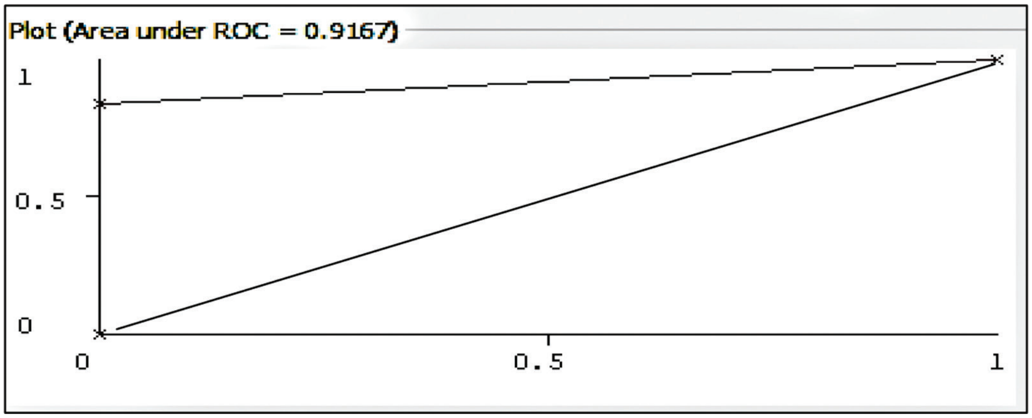
In determining the classification percentage, Table 2 shows that there are 92.78% correctly classified instances and 7.21% incorrectly classified instances this is supported by Table 3 showing the detailed accuracy by the class which the Precision weighted average of the different classification about Level of Computing is 92.4%. Moreover, this paper utilized Receiving Operating Characteristics Curve (ROC) and the Area under ROC Curve (AUC) for model accuracy a shown in Figure 4 about ROC curve and AUC curve. Also, results revealed that the class Expert has 97.9% accurate, Proficient has 94.6% accurate, Basic has 97.9% accurate, Advanced has 70.4% accurate, and Digitally Illiterate has 91.7% resulting in 94.2% model accuracy. Finally, the model has very high acceptability and accuracy in predicting the Teacher's ICT Competency.

Table 2. Summary of the evaluation of training set

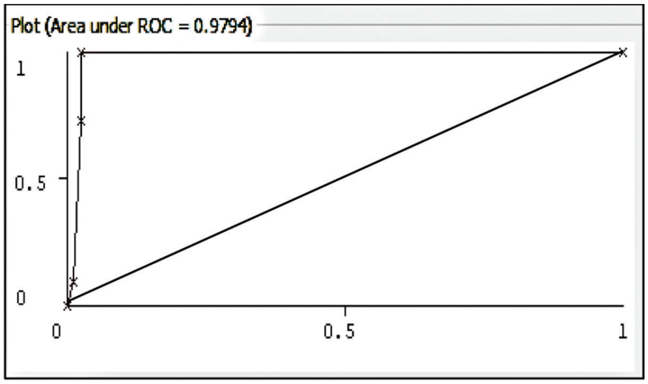
Correctly Classified Instances	90	92.7835 %
Incorrectly Classified Instances	7	7.2165 %
Kappa statistic	0.8986	
Mean absolute error	0.0401	
Root mean squared error	0.1689	
Relative absolute error	13.8624 %	
Root relative squared error	44.5156 %	
Coverage of Cases (0.95 level)	92.7835 %	
Mean rel. Region size (0.95 level)	20%	
Total Number of Instances	**97**	

Table 3. Detailed accuracy by class

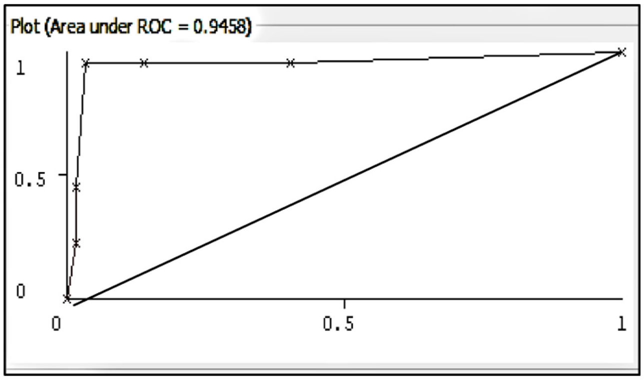
	TP Rate	FP rate	Precision	Recall	F-Measure	ROC Area	Class
	1.000	0.027	0.917	1.000	0.957	0.979	Expert
	0.950	0.035	0.950	0.950	0.950	0.946	Proficient
	1.000	0.027	0.917	1.000	0.957	0.979	Basic
	0.429	0.011	0.750	0.429	0.545	0.704	Advanced
	0.833	0.000	1.000	0.833	0.909	0.917	DigiLL
Weighted Avg.	**0.928**	**0.027**	**0.924**	**0.928**	**0.921**	**0.942**	



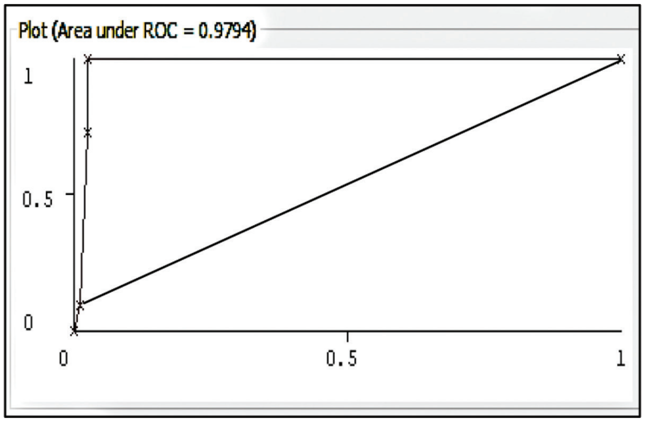
ROC Curve for Digitally Illiterate



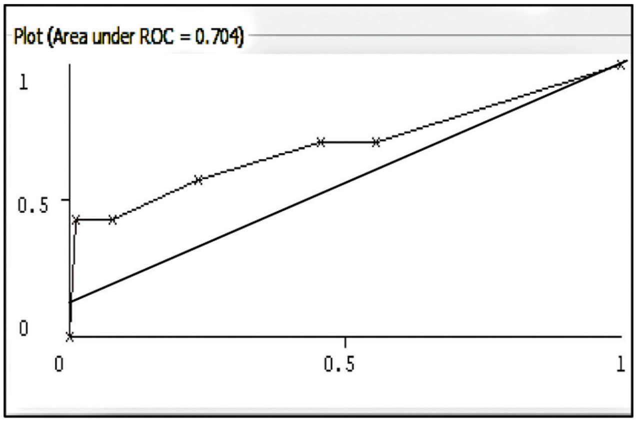
Curve for Expert



ROC Curve for Proficient



ROC Curve for Basic



ROC Curve for Advanced

Figure 4. Area under ROC curves.

5.2 Teacher's ICT Competence

ICT Competence in the field of education has shown its essential roles in the delivery of knowledge to students and teachers as well. Literature revealed that there was a definite relationship between technology acceptance and attitudes of teachers[22] who knows about using computers and related technologies which as a result, a positive impact on attitudes towards work.

The main point of this paper is to establish the capacity of the teachers in the university and create a model in using ICT Technologies in their classroom and teaching related activities. These competencies were supports to students in using ICT and ICT for instructional design and development[23]. Based on findings, there were teachers whose ICT competence were categorized as Basic which means, have basic knowledge of hardware and software, web browsers, and other multimedia devices. Moreover, some teachers were categorized as Digitally Illiterate which means, does not know how to operate computers. In the 21st century, students are far more technically inclined and more knowledgeable in using ICT tools than teachers and being categorized as Basic and Digitally Illiterate would affect the effort of the university in establishing the blended learning mode of delivery. Repositioning and capability building program relative to ICT training and seminars should be a priority to enable them to use and be productive in teaching-related activities. Frequent use of ICT in pedagogy and use of digital content must be maintained for a significant change of ICT culture in the university[24]. Finally, teachers' self-efficacy in information and communication technologies (ICT), their strategies to evaluate information, their digital competence, and use of ICT at school show a positive relationship[25].

6. Conclusion

Teacher's ICT competence in establishing a blended learning mode of delivery plays an essential role. Varied results had shown about the teacher's ICT competencies in the university which need immediate attention by sending them to training and seminars about ICT technologies for pedagogy and technology operations and concepts. Moreover, a model was established in determining the ICT competencies of the teachers with highly acceptable results as shown in the decision tree, confusion matrix, Receiving Operating Characteristics Curve (ROC), and Area under ROC Curve.

7. Recommendation

Established on the outcome of this work, it is highly commendable to send Teachers to training, workshops and further studies relative to ICT integrations in teaching and learning process. Specifically, those Teachers who were placed to be proficient, basic, and most importantly digitally illiterate as this will hamper the educational activity and learning process of both the Students and Teachers. Provide necessary ICT tools and equipment for the integration of ICT in their daily activities. Lastly, it is also encouraged to design software that will integrate and employ the developed model in order to help in decision making and strategic management of the university.

8. References

1. Rice JK. Teacher quality: Understanding the effectiveness of teacher attributes. Economic Policy Institute. 2003.
2. Croninger RG, King RJ, Rathbun, R, Masako N. Teacher qualifications and early learning Effects of certification degree and experience on first-grade student achievement, Economics of Education Review. 2007; 26(3):312–24. Crossref.
3. Garrison DR, Kanuka H. Blended learning: uncovering its transformative potential in higher education, The Internet and Higher Education. 2004; 7(2):95–105. Crossref.
4. Picciano AG, Dziuban CD, Graham CR. Blended learning: Research perspectives, Routledge. 2013; 2:1–376.
5. Masie E. The blended learning imperative. The handbook of blended learning: Global perspectives, local designs; 2006. p. 22–26.
6. Yuen AH, Hew TK. Information and Communication Technology in Educational Policies in the Asian Region.
7. Rodrigo MMT. Information and communication technology use in Philippine public and private schools, Loyola Schools Review: School of Science and Engineering. 2001; 1:122–39.
8. Marcial DE, de la Rama PA. ICT Competency level of teacher education professionals in the Central Visayas Region, Philippines, Asia Pacific Journal of Multidisciplinary Research. 2015; 3(5):28–38.
9. Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD process for extracting useful knowledge from volumes of data, Communications of the ACM. 1996; 39(11):27–34. Crossref.
10. National ICT Competency Standards for Teachers – Slide share. Date accessed: 10/6/2017. https://www.slideshare.net/.../national-ict-competency-standards-for-teachers.

11. Mohri M, Rostamizadeh A, Talwalkar A. Foundations of machine learning, MIT Press. 2012.
12. Raschka S. Predictive modeling, Supervised machine learning, and pattern classification. Machine Learning. 2014. Accessed on: 25/08/2014. http://sebastianraschka.com/Articles/2014_intro_supervised_learning.html.
13. Decision Trees Explained Using Weka, Technobium. Accessed on: 31/01/2016. http://technobium.com/decision-trees-explained-using-weka/.
14. Kotsiantis SB, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. 2007.
15. Smith L,Tansley J. Decision tree analysis. U.S. Patent Application No. 10/406,836. 2003.
16. Kaur G, Chhabra, A. Improved J48 classification algorithm for the prediction of diabetes, International Journal of Computer Applications. 2014; 98(22):1–5. Crossref.
17. Cross-Validation. Accessed on: 07/02/1997. https://www.cs.cmu.edu/~schneide/tut5/node42.html.
18. Vexler A, Yu J, Zhao Y, Hutson AD, Gurevich G. Expected p-values in light of a ROC curve analysis applied to optimal multiple testing procedures, Statistical Methods in Medical Research. 2017. Crossref.
19. Tape TG. Interpreting Diagnostic Test. Annals of Internal Medicine. University of Nebraska Medical Center; 2001.
20. Plotting and Interpreting an ROC Curve. Accessed on: 31/01/2018. http://gim.unmc.edu/dxtests/roc2.htm.
21. Baturay MH, Gökçearslan S, Ke F. The relationship among pre-service teachers computer competence attitude towards computer-assisted education and intention of technology acceptance, International Journal of Technology Enhanced Learning. 2017; 9(1):1–13. Crossref, Crossref.
22. Tondeur J, Aesaert K, Pynoo B, Braak J, Fraeyman N, Erstad O. Developing a validated instrument to measure pre-service teachers' ICT competencies: Meeting the demands of the 21st century, British Journal of Educational Technology. 2017; 48(2):462–72. Crossref.
23. Blau I, Shamir-InbalT. Digital competences and long-term ICT integration in school culture the perspective of elementary school leaders, Education and Information Technologies. 2017; 22(3):769–87. Crossref.
24. Hatlevik OE. Examining the relationship between teachers' self-efficacy, their digital competence, strategies to evaluate information, and use of ICT at school, Scandinavian Journal of Educational Research. 2017; 61(5):555–67. Crossref.
25. Fayyad U. The KDD process for extracting useful knowledge from volumes of data, Communications of the ACM. 1996; 39(11):27–34.