

TEACHERS' ICT COMPETENCY LEVEL ANALYSIS USING CLASSIFICATION ALGORITHMS

REYMON M. SANTIAÑEZ

School of Technology and Computer Studies,
Biliran Province State University Naval, Biliran, Philippines
E-mail: reymonsantianez@bipsu.edu.ph

Abstract

One of the basic requirements for providing quality teaching and learning to stakeholders is the ICT competency of teachers in the academe or institution. Individual skills are an essential aspect of achieving effective education with the aid of modern technologies. The need to support teachers' professional development is underscored by technological and pedagogical breakthroughs in ICT and Education. This study aims to predict the ICT competency level of teachers using classification algorithms, and the university used the results of the study to make interventions for teachers' need for training and development to capacitate the ICT skills of teachers. For this purpose, through WEKA data mining tool, the test options used are 10 folds cross-validation for the 7 classification algorithms: Naïve Bayes, J48, RepTree, Random Forest, Random Tree, Hoeffding Tree, and Logistic Model Tree (LMT) to predict teachers' competency level in terms of technology operations and concepts, social and ethical, pedagogical and professional. The accuracy of the algorithms was used to evaluate their performance. Furthermore, based on the overall findings of the algorithm performance, the LMT algorithm is recommended for predicting the ICT skill level of teachers. Moreover, the findings of this study will be used as a benchmark for conducting training to assist and capacitate teachers' ICT skills and gain experience in dealing with new devices, modern technologies, and new pedagogical approaches.

Keywords: Algorithm model, Classification algorithms, Data mining, Machine learning.

1. Introduction

ICT competency of teachers in the academe or institution is essential to deliver quality teaching and learning to the stakeholders. Skills for each individual are vital to achieving quality education with the help of technologies available. Technological and pedagogical developments in ICT and Education underline the importance of steadfast commitment to supporting teachers' continuous professional development [1, 2].

The challenges of daily life and work have changed dramatically in recent years due to rapid technological advancements and globalization. Many organizations and educators have noticed that success in this environment of new difficulties necessitates a new set of abilities. New solutions that demand more collaboration, flexibility, and innovation to incorporate various changing views and new technology replaced solutions that rely on fixed knowledge and linear thinking [3]. Sabado [4] specified that some complexities reinforced teachers to become flexible, adaptive, and responsive to the needs of these highly advanced and diverse digital native learners.

Higher Educational Institution (HEI) is in the flexible learning modalities to deliver quality education to the students. Still, some teachers need to be capacitated in their ICT skills to meet the demand and needs of the current situation. It is evident that technology plays an increasingly important role in society, and ICT literacy is a prerequisite for academic teachers. As a result, teachers need to build abilities that will allow them to meet the necessary competencies for learning and teaching in the twenty-first century. Despite the rising use of technology for educational reasons, there is still a perception that university teachers have a low level of ICT abilities, which has been a challenge for the institutions by Ghavifekr and Rosdy [5].

This study aims to predict the ICT competency level of teachers using classification algorithms, and the university used the result of the study to make interventions for teachers' need for training and development to capacitate the ICT skills of teachers. For this purpose the seven (7) classification algorithms were used: Naïve Bayes, J48, RepTree, Random Forest, Random Tree, Hoeffding Tree, and LMT(Logistic Model Tree) to predict teachers' competency level in terms of technology operations and concepts, social and ethical, pedagogical and professional. Classification, a supervised learning approach, can be used to predict the class of a given set of data points. Furthermore, determine the algorithm best suited for the problem in predicting the Teachers' ICT skills. The result of this study served as the basis for the training development of teachers based on the result using the classification algorithms and empower teachers to embrace 21st-century teaching techniques [6]. Furthermore, the more frequent teachers implement the ICTs, the more literate they will become. Training experiences surely generate teachers' ICT literacy [7].

2. Methods

2.1. Knowledge discovery in database (KDD) process

The study used KDD referred to as Knowledge Discovery in Database process (Fig. 1) and defined as a method of finding, transforming, and refining meaningful data and patterns from a raw database to be utilized in different domains or applications [8].

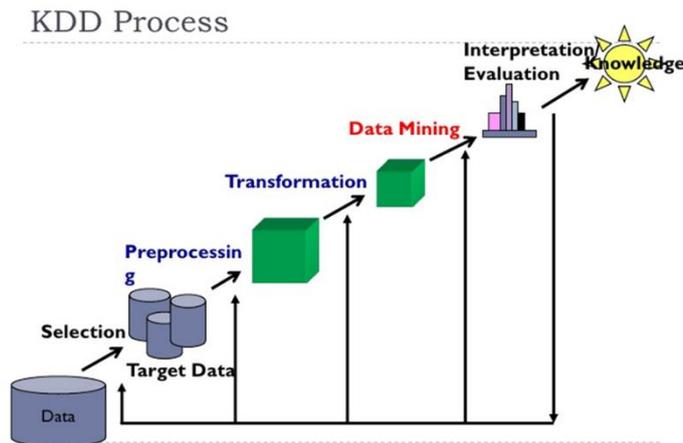


Fig. 1. KDD Process.

Data selection and integration

Datasets were collected from 225 teachers as respondents who randomly participated in the study from a different school in the university. The data was collected using Google Forms. The data was determined, selected, and segregated into meaningful sets used to evaluate for knowledge discovery. These parameters are critical for data mining because this makes as the base for it and affect what kinds of data models are formed.

Data cleaning and preprocessing

Datasets entail searching for missing data and deleting noisy, redundant, and low-quality data from the data set to increase the data accuracy and effectiveness. Predictive models for unreliable data are created to predict similarly faulty, missing, and attributional mismatched data in the future and then work it out of future processes [9].

Data transformation

This step prepares the data to be used for data mining algorithms. Hence, the data needs to be in consolidated and aggregate forms. The data is consolidated based on functions, attributes, features, etc., and uploaded in the form of a CSV file readable for the WEKA data mining tool.

Data mining

Algorithms are used to extract significant patterns from processed data, which aid in the development of prediction models. Datasets were used to identify the target labels (proficient, intermediate, novice, and basic) of the teachers. WEKA was utilized as an analytical tool in this study to discover trends from data collection using classification algorithms and techniques for building the model. The cross-validation test options were used with a ten(10) fold cross-validation. This type of test option was selected and 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 [10].

The data from the various attributes, which were responded by the respondents according to their respective levels of ICT skill, are displayed and shown in Fig. 2.

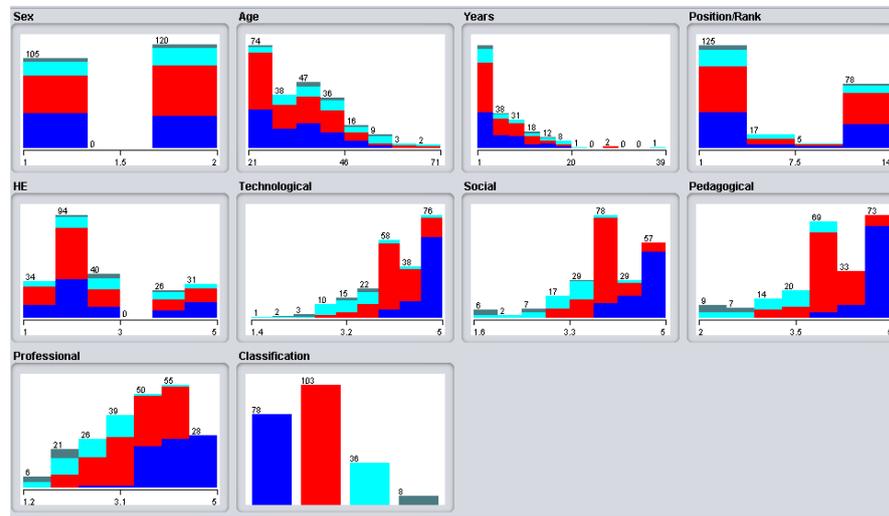


Fig. 2. Data visualization.

Interpretation and evaluation

The WEKA data mining tools were used for data analysis, and a model was constructed with various levels of accuracy based on its performance. After the trend and patterns were identified based on the different measures and have been produced using various data mining methods using classification algorithms [11]. The resulted data is interpreted and evaluated based on each algorithm's performance on teachers' ICT competency level.

2.2. Classification algorithms

The Classification algorithm is a Supervised Learning technique used to identify new observations based on training data. In classification, a program learns from the given dataset or observations and then classifies new observations into a number of classes or groups [12]. In supervised learning, models are trained using a labelled dataset, where the model learns about each type of data. Once the training process is completed, the model is tested based on test data (a subset of the training set), and then it predicts the output.

Confusion Matrix (Fig. 3) provides a matrix or table as output and describes the performance of the model. It is also known as the error matrix. The matrix consists of predictions resulting in a summarized form, which has a total number of correct predictions and incorrect predictions.

Actual	Positive	TP	FN
	Negative	FP	TN
		Positive	Negative
		Predicted	

Fig. 3. Confusion matrix.

Accuracy is one metric for evaluating classification models [13]. Informally, accuracy is the fraction of predictions our model got right. Formally, accuracy has the following definition:

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$Accuracy = \frac{T + TN}{TP + TN + FP + FN}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

2.2.1. Naïve Bayes

The Bayes Theorem-based probabilistic machine learning technique Naive Bayes is used in a wide range of categorization applications. It's not a single algorithm, but rather a group of algorithms that all follow the same principle: every pair of categorized features is independent of the others [14]. The Bayes Theorem is a simple formula that can be used to calculate conditional probabilities. Conditional probability is the chance of an event occurring given that another event has occurred (through assumption, supposition, statement, or evidence) [15, 16].

$$P(H|E) = \frac{P(E|H) * P(H)}{P(E)}$$

where $P(H|E)$ =posterior probability of the hypothesis given that the evidence is true, $P(E|H)$ =likelihood of the evidence given that the hypothesis is true, $P(H)$ =prior probability of the hypothesis, and $P(E)$ =prior probability that the evidence is true

2.2.2. J48 algorithm

Classification is the process of constructing a model of classes from a set of records having class labels. The Decision Tree Algorithm's purpose is to determine how the

attributes-vector behaves in various situations. Based on the training instances, the classes for the newly created instances are also found [17]. This algorithm generates the rules for predicting the target variable. The tree classification algorithm can be used to understand the critical distribution of data [18].

2.2.3. Random tree

The Random Tree operator works in the same way as the Decision Tree operator, with the distinction that only a random selection of characteristics is provided for each split. Using both nominal and numerical input, this operator learns decision trees. Decision trees are simple-to-understand categorization systems that are effective. The Random Tree operator, like Quinlan's C4.5 or CART, selects a random set of attributes before applying it. The size of the subset is specified by the subset ratio parameter. Tree representation of data has the advantage of being intelligible and easy to comprehend when compared to other methods. The goal is to develop a classification model that predicts label value from numerous dataset input attributes. Each input attribute corresponds to one of the internal nodes of the tree. The number of available input attribute values is the same as the number of internal node edges. Each leaf node represents a label value based on the values of the input attributes provided by the path from root to leaf [19].

2.2.4. Random forest

This method is made up of many separate decision trees that work together to form an ensemble. The random forest (Fig. 4) produces a class prediction for each tree, and the class with the highest votes becomes our model prediction. The random forest is a classification algorithm that uses numerous decision trees to classify data. When developing each tree, it uses bagging and feature randomization to create an uncorrelated forest of trees whose committee prediction is more accurate than any individual tree [20].

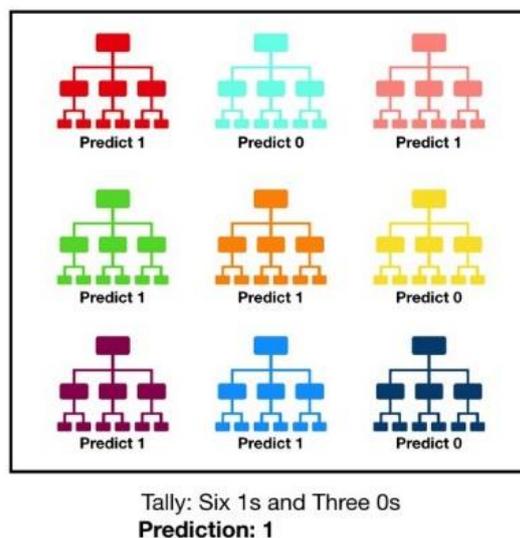


Fig. 4. Random forest.

2.2.5. REPTree

This is a fast decision tree learner that uses the C4.5 algorithm to build classification (discrete outcome) or regression trees (continuous outcome). It uses information gain/variance to build a regression/decision tree and prunes it using reduced-error pruning (with back-fitting) [21].

2.2.6. Logistic model tree

The logistic model tree (LMT), a new approach that combines classic decision tree (DT) induction and the linear logistic regression algorithm into a single tree, has recently been proposed as a replacement for DT-based learning algorithms [22]. Furthermore, this classification algorithm is an associated supervised training algorithm that combines logistic regression (LR) and decision tree learning.

2.2.7. Hoeffding tree

A Hoeffding tree (VFDT) is a decision tree induction system that can learn from enormous data streams incrementally and at any time, given that the distribution generating samples do not vary over time. Hoeffding trees take advantage of the notion that a small sample size can often be sufficient for determining the best splitting attribute. The assumption is statistically supported by the Hoeffding bound, which estimates the number of observations required to estimate given statistics within a specified precision [23].

3. Results and Discussion

3.1. Naive Bayes.

Table 1 shows the performance of the Naïve Bayes algorithm performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance also shows the correctly classified instances of 195 equivalent to 86.67%, and 30 or 13.33% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.793, TP Rate 0.867, Precision is 0.871, Recall is 0.867, ROC Area is 0.974, and PRC Area is 0.943, respectively.

Table 1. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.793	0.867	0.871	0.867	0.974	0.943	195 86.67%	30 13.33%

Figure 5 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of the Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 71 Advanced objects as Advanced, 2 as an Intermediate, 0 as a Novice, and 0 as a Basic leading to 4 misclassifications.
- The algorithm has classified 92 Intermediate objects as Intermediate, 7 as Advanced, 7 as Novice, and 0 as Basic, leading to 17 misclassifications.

- The algorithm has classified 26 Novice objects as Novice, 0 Advanced, 9 Intermediate, and 2 Basic leading to 11 misclassifications.
- The algorithm has classified 6 Basic objects as Basic, 0 Advanced, 0 Intermediate, and 3 Novice leading to 3 misclassifications.

True Label	Advanced	71	7	0	0
	Intermediate	2	92	9	0
	Novice	0	7	26	3
	Basic	0	0	2	6
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 5. Confusion matrix (Naïve Bayes algorithm).

3.2. J48 algorithm

Table 2 shows the J48 algorithm's performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance, also shows the correctly classified instances of 174 equivalent to 77.33%, and 51 or 22.67% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.6484, TP Rate 0.773, Precision is 0.788, Recall is 0.773, ROC Area is 0.853, and PRC Area is 0.714, respectively.

Table 2. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.6484	0.773	0.788	0.773	0.853	0.714	174 77.33%	51 22.67%

Figure 6 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 65 Advanced objects as Advanced, 8 as an Intermediate, 0 as a Novice, and 0 as a Basic leading to 8 misclassifications.
- The algorithm has classified 83 Intermediate objects as Intermediate, 13 as Advanced, 10 as Novice, and 0 as Basic, leading to nine (23) misclassifications.
- The algorithm has classified 21 Novice objects as Novice, 0 Advanced, 12 Intermediate, and 3 Basic leading to 15 misclassifications.

- The algorithm has classified 5 Basic objects as Basic, 0 Advanced, 0 Intermediate, and five (5) Novice leading to 5 misclassifications.

True Label	Advanced	65	13	0	0
	Intermediate	8	83	12	0
	Novice	0	10	21	5
	Basic	0	0	3	5
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 6. Confusion matrix (J48 algorithm).

3.3. Random forest

Table 3 shows the Random Forest algorithm's performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance also shows the correctly classified instances of 194 equivalent to 86.22%, and 31 or 13.78% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.7814, TP Rate 0.862, Precision is 0.858, Recall is 0.862, ROC Area is 0.971, and PRC Area is 0.917, respectively.

Table 3. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.7814	0.862	0.858	0.862	0.971	0.917	194 86.22%	31 13.78%

Figure 7 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of the Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 71 Advanced objects as Advanced, 4 as Intermediate, 0 as a Novice, and 0 as a Basic leading to 4 misclassifications.
- The algorithm has classified 97 Intermediate objects as Intermediate, 7 as Advanced, 10 as Novice, and 0 as Basic, leading to 17 misclassifications.
- The algorithm has classified 23 Novice objects as Novice, 0 Advanced, 2 Intermediate, and 5 Basic leading to 7 misclassifications.

- The algorithm has classified 3 Basic objects as Basic, 0 Advanced, 0 Intermediate, and 3 Novice leading to 3 misclassifications.

True Label	Advanced	71	7	0	0
	Intermediate	4	97	2	0
	Novice	0	10	23	3
	Basic	0	0	5	3
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 7. Confusion matrix (Random forest algorithm).

3.4. REPTree

Table 4 shows the Random Forest algorithm's performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance also shows the correctly classified instances of 170 equivalent to 75.56% and 55 or 24.44% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.6121, TP Rate 0.756, Precision is 0.754, Recall is 0.756, ROC Area is 0.854, and PRC Area is 0.715, respectively.

Table 4. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.6121	0.756	0.754	0.756	0.854	0.715	170	75.56%
						55	24.44%

Figure 8 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of the Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 64 Advanced objects as Advanced, 13 as Intermediate, 0 as Novice, and 0 as a Basic leading to 13 misclassifications.
- The algorithm has classified 83 Intermediate objects as Intermediate, 14 as Advanced, 15 as Novice, and 5 as Basic, leading to 34 misclassifications.
- The algorithm has classified 20 Novice objects as Basic, 0 as Advance, 7 as Intermediate, and 5 Basic leading to 12 misclassifications.
- The algorithm has classified 3 Basic objects as Basic, 0 as Advance, 0 as Intermediate, and 1 Novice leading to 1 misclassification.

True Label	Advanced	64	14	0	0
	Intermediate	13	83	7	0
	Novice	0	15	20	1
	Basic	0	0	5	3
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 8. Confusion matrix (REP tree algorithm).

3.5. Random tree

Table 5 shows the Random Forest algorithm's performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance also shows the correctly classified instances of 168 equivalent to 74.67%, and 57 or 25.33% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.6082, TP Rate 0.747, Precision is 0.745, Recall is 0.747, ROC Area is 0.810, and PRC Area is 0.653, respectively.

Table 5. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.6082	0.747	0.745	0.747	0.810	0.653	168 74.67%	57 25.33%

Figure 9 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 69 Advanced objects as Advanced, 15 as an Intermediate, 0 as Novice, and 0 as a Basic leading to 15 misclassifications.
- The algorithm has classified 76 Intermediate objects as Intermediate, 9 as Advanced, 11 as Novice, and 0 as Basic, leading to 20 misclassifications.
- The algorithm has classified 21 Novice objects as Basic, 0 as Advance, 12 as Intermediate, and 6 Basic leading to 18 misclassifications.
- The algorithm has classified 2 Basic objects as Basic, 0 as Advance, 0 as Intermediate, and 4 Novice leading to 4 misclassifications.

True Label	Advanced	69	9	0	0
	Intermediate	15	76	12	0
	Novice	0	11	21	4
	Basic	0	0	6	2
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 9. Confusion matrix (Random tree algorithm).

3.6. LMT (Logistic model tree)

Table 6 shows the LMT (Logistic Model Tree) algorithm's performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance also shows the correctly classified instances of 211 equivalent to 93.78% and 14 or 6.22% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.9022, TP Rate 0.938, Precision is 0.939, Recall is 0.938, ROC Area is 0.988, and PRC Area is 0.967, respectively.

Table 6. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.9022	0.938	0.939	0.938	0.988	0.967	211 93.78%	14 6.22%

Figure 10 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 74 Advanced objects as Advanced, 0 as an Intermediate, 0 as Novice, and 0 as a Basic leading to 0 misclassifications.
- The algorithm has classified 102 Intermediate objects as Intermediate, 4 as Advanced, 5 as Novice, and 0 as Basic, leading to 9 misclassifications.
- The algorithm has classified 29 Novice objects as Basic, 0 as Advance, 1 as Intermediate, and 2 Basic leading to 3 misclassifications.
- The algorithm has classified 6 Basic objects as Basic, 0 as Advance, 0 as Intermediate, and 2 Novice leading to 2 misclassifications.

True Label	Advanced	74	4	0	0
	Intermediate	0	102	1	0
	Novice	0	5	29	2
	Basic	0	0	2	6
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 10. Confusion matrix (LMT algorithm).

3.7. Hoeffding Tree

Table 7 shows the Hoeffding Tree algorithm's performance in terms of accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, the algorithm performance also shows the correctly classified instances of 193 equivalent to 85.78%, and 32 or 14.22% incorrectly classified from the datasets. The classes from Kappa Statistic with the value of 0.7789, TP Rate 0.858, Precision is 0.863, Recall is 0.858, ROC Area is 0.973, and PRC Area is 0.939, respectively.

Table 7. Performance of the algorithm.

Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area	Correctly Classified Instances	Incorrectly Classified Instances
0.7789	0.858	0.863	0.858	0.973	0.939	193	85.78%
						32	14.22%

Figure 11 shows the confusion matrix and illustrates the correctly classified instances and the incorrectly classified instances of Teachers' ICT Competency Level. Likewise, the Confusion Matrix concluded as:

- The algorithm has classified 69 Advanced objects as Advanced, 2 as an Intermediate, 0 as Novice, and 0 as a Basic leading to 2 misclassifications.
- The algorithm has classified 92 Intermediate objects as Intermediate, 9 as Advanced, 7 as Novice, and 0 as Basic, leading to 16 misclassifications.
- The algorithm has classified 26 Novice objects as Basic, 0 as Advance, 9 as Intermediate, and 2 Basic leading to 11 misclassifications.
- The algorithm has classified 6 Basic objects as Basic, 0 as Advance, 0 as Intermediate, and 3 Novice leading to 3 misclassifications.

True Label	Advanced	69	9	0	0
	Intermediate	2	92	9	0
	Novice	0	7	26	3
	Basic	0	0	2	6
		Advanced	Intermediate	Novice	Basic
		Predicted Label			

Fig. 11. Confusion matrix (Hoeffding tree algorithm).

3.8. Performance comparison between different classification algorithms

To ensure the validity and accuracy of the model, the results are compared based on its classification algorithm's performance [24, 25].

Table 8 shows the accuracy summary of the evaluation with the 7 Classification Algorithm, in Naïve Bayes 86.6667% as correctly classified instances and 13.3333% incorrectly classified instances with 0.01 seconds execution time, in J48 77.3333% as correctly classified instances and 22.6667% incorrectly classified instances with 0.04 seconds execution time in Random Forest 86.2222% as correctly classified instances and 13.7778% incorrectly classified instances with 0.18 seconds execution time and in REPTree 75.5556% as correctly classified instances and 24.4444% incorrectly classified instances with 0.01 seconds execution time in RandomTree 74.6667% as correctly classified instances and 25.3333% incorrectly classified instances with 0 seconds execution time and LMT 93.7778% as correctly classified instances and 6.22% incorrectly classified instances with 0.6 seconds execution time and in Hoeffding Tree 85.78% correctly classified instances and 14.22% incorrectly classified instances with 0.04 seconds execution time . This indicates that the model is acceptable in terms of accuracy in predicting the Teachers' ICT Competency using LMT (Logistic Model Tree) algorithm with the highest correctly classified instances of 93.7778%.

Table 9 shows the weighted average values of the classification algorithms in terms of summary and accuracy by class of Kappa Statistic, TP Rate, Precision, Recall, ROC Area, and PRC Area. Moreover, on the performance of the algorithms, it shows that LMT has the highest values in all of the classes from Kappa Statistic with the value of 0.9022, TP Rate 0.938, Precision is 0.939, ROC Area is 0.988, and PRC Area which is 0.967 respectively. On the other side, opposite to the LMT is Random Tree having a performance value of 0.6082 on Kappa Statistic, TP Rate which is 0.747, Precision 0.745, Recall 0.747, ROC Area is 0.810, and PRC Area with the value of 0.653.

Table 8. Classifiers accuracy summary.

Algorithm	Time taken to build the model	Correctly Classified Instances	Incorrectly Classified Instances
<i>Naïve Bayes</i>	0.01	195 - 86.67 %	30 - 13.33 %
<i>J48</i>	0.04	174 - 77.33%	51 - 22.67 %
<i>Random Forest</i>	0.18	194 - 86.22 %	31 - 13.78 %
<i>REPTree</i>	0.01	170 - 75.56 %	55 - 24.44 %
<i>Random Tree</i>	0	168 - 74.67 %	57 - 25.33 %
<i>LMT</i>	0.6	211 93.78 %	14 - 6.22 %
<i>Hoeffding Tree</i>	0.04	193 85.78 %	32 - 14.22 %

Table 9. Performance of the algorithm's summary.

Algorithms	Kappa Statistic	TP Rate	Precision	Recall	ROC Area	PRC Area
<i>Naïve Bayes</i>	0.793	0.867	0.871	0.867	0.974	0.943
<i>J48</i>	0.6484	0.773	0.788	0.773	0.853	0.714
<i>Random Forest</i>	0.7814	0.862	0.858	0.862	0.971	0.917
<i>REPTree</i>	0.6121	0.756	0.754	0.756	0.854	0.715
<i>Random Tree</i>	0.6082	0.747	0.745	0.747	0.810	0.653
<i>LMT</i>	0.9022	0.938	0.939	0.938	0.988	0.967
<i>Hoeffding Tree</i>	0.7789	0.858	0.863	0.858	0.973	0.939

4. Conclusions

Results from the different classification algorithms identified based on the performance as consistently best suited for predicting the ICT competency level of teachers in the university. The results also identified that in all of the classification algorithms used to predict the ICT competency of teachers for the creation of the model the Logistic Model Tree (LMT) has the highest accuracy and the less performance on its accuracy is the Random Tree algorithm. Therefore, the algorithm with a high accuracy level and served as a model is recommended to be used for the prediction of teachers' ICT competency levels. Furthermore, the results of this study will be served as a benchmark to conduct a training program that would help and capacitate the ICT skills of teachers [26].

References

1. Thakral, P. (2015). Role of ICT in professional development of teachers. *Learning Community*, 6(1), 127-133.
2. UNESCO (2018). UNESCO ICT competency framework for teachers (ver.3). United Nations Educational, Scientific and Cultural Organization, 66.

3. ACTRC (2015). 21ST Century Skills in the Philippines. Retrieved April 20, 2021, from <https://actrc.org/projects/current-projects/21st-century-skills-in-the-philippines/>
4. Sabado, A. C. (2016). Education in 21'st Century Philippines. Retrieved May 15, 2021, from <https://www.pressreader.com/philippines/sunstar-pampanga/20160619/281784218382642>
5. Ghavifekr, S.; and Rosdy, W.A.W. (2015). Teaching and learning with technology: effectiveness of ICT integration in school. *International Journal of Research in Education and Science (IJRES)*, 1(2), 175-191.
6. Bingimlas, K.A. (2009). Barriers to the successful integration of ICT in teaching and learning environments: A review of the literature. *Euroasia Journal of Mathematics, Science and Technology Education*, 5(3), 235-245.
7. Hafifah, G.N.; and Sulisty, G.H. (2020). Teachers' ICT literacy and ICT integration in ELT in Indonesian higher education setting. *Turkish Online Journal of Distance Education*, 21(3), 186-198.
8. Sharma, R. (2020). KDD process in data mining: What you need to know? Retrieved May 20, 2021, from <https://www.upgrad.com/blog/kdd-process-data-mining/>
9. Hotz, N. (2021). KDD and data mining. Retrieved August 15, 2021, from <https://www.datascience-pm.com/kdd-and-data-mining/>
10. Gupta, P. (2017). Cross-validation in machine learning, towards data science. Retrieved May 15, 2021, from <https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f>
11. Valadkhani, M. (2016). Knowledge discovery in data (KDD) process. Retrieved May 16, 2021, from <https://www.linkedin.com/pulse/knowledge-discovery-data-kdd-process-mohammad-valadkhani>
12. Javatpoint (2021). Classification algorithm in machine learning. Retrieved April 16, 2021, from <https://www.javatpoint.com/classification-algorithm-in-machine-learning>
13. Google. Classification: Accuracy. Retrieved May 16, 2021, from <https://developers.google.com/machine-learning/crash-course/classification/accuracy>
14. Kumar, N. (2020). Naive Bayes classifiers. Retrieved May 16, 2021, from <https://www.geeksforgeeks.org/naive-bayes-classifiers/>
15. Chauhan, N.S. (2020). Naïve Bayes algorithm: Everything you need to know, KD Nuggets. Retrieved May 16, 2021, from <https://www.kdnuggets.com/2020/06/naive-bayes-algorithm-everything.html>
16. Gamal, B. (2020). Naïve Bayes algorithm. Retrieved April 5, 2021, from <https://medium.com/analytics-vidhya/na%C3%AFve-bayes-algorithm-5bf31e9032a2>
17. Korting, T.S. (2006). C4. 5 algorithm and multivariate decision trees. Image Processing Division, National Institute for Space Research – INPE.
18. Nadali, A.; Kakhky, E.N.; and Nosratabadi, H.E. (2011). Evaluating the success level of data mining projects based on CRISP-DM methodology by a fuzzy expert system. *Proceedings of the 2011 3rd International Conference on Electronics Computer Technology*, Kanyakumari, India, 161-165.

19. Rapidminer (2022). Random tree. Retrieved May 20, 2021, from https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/trees/random_tree.html
20. Yiu, T. (2019). Understanding random forest. Retrieved May 15, 2021, from <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>
21. Al Snousy, M.B.; El-Deeb, H.M.; Badran, K.; and Al Khilil, I.A. (2011). Suite of decision tree-based classification algorithms on cancer gene expression data. *Egyptian Informatics Journal*, 12(2), 73-82.
22. Colkesen, I.; and Kavzoglu, T. (2017). The use of logistic model tree (LMT) for pixel- and object-based classifications using high resolution WorldView-2 imagery. *Geocarto International*, 32(1), 71-86.
23. Weka. Class Hoeffding Tree. Retrieved May 16, 2021, from <https://weka.sourceforge.io/doc.dev/weka/classifiers/trees/HoeffdingTree.html>
24. Talingdan, J.A. (2019). Performance comparison of different classification algorithms for household poverty classification. *Proceedings of the 2019 4th International Conference on Information Systems Engineering (ICISE)*, Shanghai, China, 11-15.
25. Çiğşar, B.; and Ünal, D. (2019). Comparison of data mining classification algorithms determining the default risk. *Scientific Programming*, Volume 2019, Article ID 8706505.
26. Apolinar-Gotardo, M. (2020). A classification algorithm analysis of students' ICT competency level using data mining technique. *International Journal of Scientific & Technology Research*, 9(3), 3256-3258.