Student Performance Prediction on Internet Mediated Environments using Decision Trees

Esther Khakata
Strathmore University
Faculty of Information Technology
Nairobi, Kenya

Vincent Omwenga
Strathmore University
Faculty of Information Technology
Nairobi, Kenya

Simon Msanjila
Mzumbe University
Faculty of Science and Technology
Dar es Salaam, Tanzania

ABSTRACT

A decision tree is a data classification technique that is used in mining data. The use of data mining techniques on data related to student performance assists in extracting valuable information from large data sets available within the institutions. This paper discusses the factors that influence the student learning process in an internet mediated environment and presents results from data collected within universities in Kenya. Additionally, this paper uses a decision tree to affirm these factors that affect student performance within the universities. The decision tree generated acts as a prediction tool that assists in the prediction of student performance while using internet technology in the learning process. This enables us determine whether a student is likely to perform well or not while using internet technology in the learning process.

General Terms

Student, Performance, Prediction, Decision Tree

Keywords

Student performance, data mining, decision trees, performance prediction

1. INTRODUCTION

Predicting student performance on internet mediated platforms has been greatly influenced by the enormous quantity of data that is available online which spans hundreds of petabytes [1]. This is mainly attributed to the growth from Web 1.0 services to Web 2.0 services where data is available on many platforms [2].

In terms of learning on internet mediated platforms, research shows that students have been engaging with online course materials by communicating with online communities on forums, asking questions and receiving appropriate feedback whenever necessary [3]. This data about the different learning activities can be collected, profiled and used to predict the performance of the learners while they use internet mediated platforms in their learning process [2].

Learning in an internetworking environment is influenced by a number of factors specific to institutions and students. In this paper, an internet mediated environment is defined as an environment where a range of daily activities and processes are conducted online by the people around the setting through the use of internet technology. This technology has superseded existing means of communication by overpowering the obstacles of distance, time and personal schedules. Consequently, ease of learning has been seen in the use of readily available communication channels, internet-aided learning and instruction, availability of digital books along with other valuable resources essential for learning [4].

2. FACTORS AFFECTING LEARNING ON INTERNET MEDIATED ENVIRONMENTS

There exists a number of factors that influence how the learning process takes place on internet mediated environments. These factors include the knowledge and competence levels of the learners in using internet technology, their effort in using the technology, the physical learning environment, the attitude and behavior of the learners towards the technology, belief in the importance of the technology and the investment costs required to access internet technology.

Data on each of the above factors was collected from universities by the use of questionnaire surveys. The survey method involved the issuance of 1,100 copies to students within 12 public universities and 8 private universities in Nairobi, Kenya. The survey yielded 796 questionnaires out of which 747 were usable, with a response rate of 72%. The data was then analyzed and the results were used also used in the generation of the decision tree used in student performance prediction.

2.1 Knowledge and Competence Levels

The knowledge and competence level of students in universities as they interact with internet technology improves over time. Though some students lack the simplest technological knowledge and skills to use the internet in the learning process, this is seen to gradually improve with their years in the university [5]. Whereas some students find the use of internet in the learning process an enjoyable task, those with insufficient skill sets find the use of the technology a daunting task due to the large sea of information available to them within an instance. Students who do not possess the classical information-seeking behavior do not benefit from the available modern information technologies. This situation can be combated by offering adequate training opportunities to the students in these learning institutions in order to ensure that they are equipped with the necessary skills they need to successfully use the internet in their learning process. Due to lack of searching skills among students, internet technology remains an obstacle to the effective use of the internet in the learning process [6]. Despite the challenges undergone by students in the use of internet facilities, it is obvious that the use of the resource enhances their academic performances in various fields. They therefore need to overlook these existing challenges and embrace the numerous benefits offered by the technological facility to enrich their intellects with the available knowledge.

Examining the importance of internet technology in the learning process, the data collected revealed that many respondents considered the use of internet technology a vital component in their learning process. The respondents were required to choose one option indicated on a 5-point scale

whether they considered the technology relevant to their studies or not. The 5-point scale included irrelevant, slightly relevant, not sure, relevant and very relevant. The number of respondents is given by **n.**

Table 1. Importance of Technology

How relevant has the internet been in your studies?				
Irrelevant (n=8)	Slightly relevant (n=30)	Not sure (n=19)	Relevant (n=305)	Very relevant (n=382)
1.07%	4.02%	2.54%	40.83%	51.34%
5.09%		2.54%	92.17%	

Table 1 shows that only 5.09% of the respondents did not consider the use of internet technology relevant in their learning process, suggesting that they did not use the technology in learning and they did not consider the technology as a useful resource in the learning process as well. However, 92.17% of the respondents considered internet technology relevant in their learning process, meaning that, they used the technology in their studies and they found the technology significantly important within their learning environment.

In considering the different purposes for which internet technology was used by students and how often it was used for learning purposes, the respondents were allowed to select the different ways in which they were using internet technology in their learning. They were given six choices and they were free to select more than one option. The choices were research work, assignments, communication, music/movies, meeting people/friends/peers and games.

Table 2. Purpose of internet technology

Please indicate for what purpose you use the internet					
Research work (n=605)	Assignments (n=575)	Communication (n=406)	Music/movies (n=236)	Meeting people/friends/peers (n=143)	Games (n=108)
81%	77%	54%	32%	19%	14%

Table 2 revealed that, indeed, respondents used internet technology for a number of tasks in their learning process. For instance, 81% of the students used internet technology for their research work; 77% of the respondents used the technology for their class assignments; 54% used the internet for communication purposes; 32% of the respondents used the internet for downloading music and movies; 19% used the resource for meeting people and 14% used the internet for gaming purposes. Hence, these statistics show that the students were using the technology greatly for learning purposes as compared to other leisure related activities.

2.2 Student Effort

Student effort refers to the amount of time and energy expended by students in order to achieve a previously set objective [7]. Student effort is a valuable predictor of success when considering cases of individual learners. It marks how motivated a student is in the subject area or how discouraged they are in the learning process [8]. Effort indicates how engaged the learners are in their academic work, whether they

keep trying hard, they continuously work hard and whether they pay attention to their studies by constancy in their academic work [9] [10]. Effort can be easily controlled and changed by students voluntarily. Since individual effort affects the learning process and academic performance, effort influences the perceptions that individuals have about their capabilities and the technology they need to use in the learning process [10].

While examining student perceptions the Technology Acceptance model (TAM) posits that the students need to find the technology being useful (Perceived Usefulness - PU) and easy to use (Perceived Ease of Use - PEOU). Table 3 shows the results obtained from the collected data.

Table 3. PU and PEOU

Perceived Usefulness	n	Disagree	Neutral	Agree
The internet allows me to increase my productivity in my studies	747	4.2%	9.8%	86.1%
The internet has led me to rely less on hard copy text for my studies	747	10.3%	19.1%	70.5%
The internet has enhanced the quality of the work I do in my studies	747	3.5%	17.8%	78.7%
The internet gives me a great sense of accomplishment after using it for learning purposes	747	4.4%	18.5%	77.1%
The internet has enhanced my performance in my studies	747	4.7%	16.5%	78.8%
I find internet useful in my studies	747	2.4%	9.5%	88.1%
Using the internet enables me to accomplish tasks more quickly	747	3.1%	12.4%	84.5%
I feel that using internet resources gives me a great deal of opportunity for my studies	747	3.7%	15.8%	80.5%
I find that I have fewer challenges with my studies than my course mates/friends due to use of the internet	747	10.4%	25.3%	64.3%
I can competently complete any assigned task using internet	747	4.7%	15.8%	79.7%
Perceived Ease of Use	n	Disagree	Neutral	Agree
I find the internet easy to use and enjoyable	747	2.9%	12.3%	84.7%
Using the internet is a pleasant experience for me	747	3.6%	15%	81.4%
The internet is a necessary tool in my academics	747	4.8%	11.4%	83.8%

From the results given on table 3, it can be seen that the respondents found internet technology useful in their learning process since they managed to accomplish tasks by using the technology. Additionally, the respondents found the technology easy to use and enjoyable and hence embraced its use in their learning process.

2.3 Investment in the Physical Learning Environment

In considering the physical learning environment of a student, research has noted that the student's achievement is greatly affected by the learning environment [11]. Essentially, the learning environment refers to all the components and undertakings within which learning takes place. Undeniably, the learning environment plays a crucial role in shaping the quality of achievements in the learning environment [12]. Researchers have concurred with this position by alluding to the fact that, for a student to learn well, there is need to have an enabling environment that is free from possible intellectual disturbances. This therefore implies that, the condition of the learner should be constantly checked to give conducive learning environs [13] [14] [15]. The environment needs to focus on providing information that leads to the mastery of a skill, and not passive information [16]. Additionally, teachers in these environment need to assist students by helping them stimulate their knowledge creation skills.

In order to establish whether the respondents' physical environment was suitable for accessing internet technology in the learning environment, there was need to check whether there were designated places where students accessed the resource from and whether the students were supported in using the resource within the university. Hence, respondents were presented with a set of 15 items on a 5-point Likert scale that ranged between 1 for 'Never' to 5 for 'Always'. A higher score on the items, that is > 3 was associated with a better physical environment that supported access to internet technology while lower levels on the items given, that is, < 3, was associated with poor physical environment that did not support access to internet technology in the learning process.

Table 4. Physical environment and general internet infrastructure

Physical environment and general internet infrastructure	n	Mean	Standard Deviation
Spacious computer labs are available with adequate light, controlled temperatures and minimal noises.	747	3.62	1.22
There are many adequately equipped labs that are available with internet connectivity.	747	3.55	1.25
Internet connectivity is available all around the university.	747	3.25	1.38
I have a personal laptop for internet access.	747	3.61	1.39
I have a smart gadget(s) that I use for internet access in lecture halls.	747	3.93	1.24
Available study areas are comfortable (clean, well-organized) and with internet connectivity.	747	3.78	1.17
Lecturers have computers with internet connectivity.	747	3.63	1.24
I have access to consultation rooms with internet connection to enable me meet my lecturers.	747	3.04	1.41
I use the e-learning portal to get access to my learning materials.	747	3.59	1.35
The lecture theatres have internet connectivity and course work materials can be viewed in class using LCD projectors.	747	3.27	1.40
Adequate assistance is offered to students to ensure that they can access the internet in the university.	747	3.37	1.32
University hostels/student hostels around the university have adequate internet access.	747	2.81	1.40
Internet access is open and free to all within the university.	747	3.56	1.37
The university library has internet access for all library users.	747	4.04	1.16
Research materials are accessible from different online databases.	747	4.06	1.12

The data collected as seen in table 4 revealed that the physical environment and general internet infrastructure within the universities was occasionally meeting the needs of the learners within their institutions. It was also confirmed that many institutions often provided accessibility to internet technology within their library premise. This parameter scored a high mean of 4.04 (standard deviation = 1.16), implying that accessibility to internet was mainly confined to the library premises. At the same time, the students also affirmed that they could be in a position to access different online databases while within the library premise. This parameter scored the highest mean of 4.06 (standard deviation = 1.12).

On the other hand, it was evident that the university hostels rarely had accessibility to internet technology. This parameter scored the lowest mean of 2.81 (standard deviation = 1.40), inferring that the institutions had not invested in the provision of the resource to the students in their halls of residence. Similarly, the universities seem not to have consultation rooms with internet connectivity to enable students meet with their lecturers. This parameter scored a mean of 3.04 (standard deviation = 1.41), implying that access to such kind of facilities within the institutions was a rare occurrence.

2.4 Attitude and Behavior

According to the Theory of Reasoned Action (TRA) [17], the attitude of a student towards a technology is a characteristic of their individual behavior. A person's attitude towards performing a specific behavior refers to their individual views of personal desirability to perform the behavior [18] [19]. Attitude depends on the prospects about effects of outcomes that result from their behavior. Hence, the attitude of an individual is the result of the evaluation of their behavior and the possible outcomes [17]. Though attitude has been defined as the evaluation of a person's desire to use a system [20], this paper defines attitude as the students evaluation of the consequences related to the use of internet technology. Student attitude in the use of internet technology in the learning process is greatly fostered by having an internetworked learning institution [21]. The use of internet

technology in teaching and learning encourages students to use the technology since the students find themselves in an internet-enriched learning environment. This positively affects the attitude of the students in using the technology for learning purposes [22]. Students who embrace the learning process with positive attitude end up using internet technology in their learning process with the same positive attitude [23].

In examining students' attitude, this aspect tested the approach embraced by a student while using the technology in learning. Specifically, this measure was operationalized by examining student confidence in using the technology, whether the students were interested in what the technology offered for their studies and whether the student found the online resources necessary in their academics. It also checked whether the student used the materials available as the only source of updated information and whether they used learning systems in their studies.

The measure contained 12 items. The respondents were required to indicate the extent to which the items given were true regarding their attitude towards the technology. This measure presented items on a 5-point Likert scale ranging from strongly agree (5) to strongly disagree (1).

Table 5. Student attitude

Student Attitude	n	Mean	Standard Deviation
I am afraid that if I begin to use internet I will become dependent upon it and lose some of my reasoning skills	747	4.19	0.79
I am sure that with time and practice I will be as comfortable working with internet as I am in working with a library of books	747	3.90	0.96
I keep up with the advances happening in the internet field	747	4.31	0.73
I dislike working with internet since it appears like a machine that is smarter than I am	747	4.35	0.71
I have difficulty in understanding the technical aspects of internet	747	4.43	0.76
I hesitate to use internet for fear of getting too much information that requires a lot of sifting	747	2.87	1.29
I have avoided internet because it is unfamiliar and somewhat intimidating to me	747	4.03	0.93
I seek information from the internet for learning activities, for example, assignments and projects	747	3.87	0.96
I search for materials from the internet to complete my assignments and projects	747	2.27	1.29
I use the Internet as the main source of information for my studies	747	2.54	1.23
I use the internet to access the Learning Management System/E-learning portal as part of my	747	2.39	1.23
learning activity			
I seek the latest information online to enhance my knowledge related to the courses taken in the university	747	2.16	1.27
I use internet forums to exchange opinions on academic matters with my friends	747	3.73	1.05

While considering student attitude towards the use of internet technology in the learning process as shown in table 5, the parameter that scored the highest was, 'I have difficulty in understanding the technical aspects of internet'. This parameter scored the highest mean of 4.43 (standard deviation = 0.76), implying that majority of the respondents had a challenge with the technical aspects of the internet and they could not figure out its complex technical aspects, for instance, its infrastructure.

Similarly, another parameter that scored a high mean was 'I dislike working with internet since it appears like a machine that is smarter than I am'. This parameter scored a mean of 4.35 (standard deviation = 0.71), denoting that many respondents found the internet smarter than them and ended up not wanting to work with it. On the other hand, the lowest mean was obtained from 'I seek the latest information online to enhance knowledge to courses taken at the university'. This parameter scored the lowest mean of 2.16 (standard deviation = 1.27), suggesting that the respondents were not fond of using the internet to get the latest information about their fields of their study.

2.5 Belief in the Importance of Technology

The belief in the importance of a technology aims at showing the technical developments that have so far been achieved over time and the general feel of the current users of the same technology [24]. The development of the technology shows itself in the predictable and traceable path that can be seen so far in its use. Specifically, in its ease of use and its perceived usefulness [25]. This is also seen in the effects produced by the technology in solving current existing challenges and supporting further development of the work for which the technology engages in [26]. This is defined as the Task technology Fit.

Considering the Task Technology Fit (TTF) Model, this model takes into consideration the task to be accomplished and the technology available. In a situation where the technology available was used in undertaking a task, then, the technology was found to be useful and the individual performance improved. In this context, the task to be undertaken was the learning process and the technology to be used was internet technology. The use of internet technology in the learning process was examined by checking the different ways a user needed to use the technology in order to perform better as shown in table 6.

Table 6. Task-Technology Fit

Tasks-internet technology fit	n	Irrelevant	Not Sure	Relevant
Source of updated academic information	747	6.4%	6.6%	87%
Access to learning materials through an e-learning portal/learning management system or to upload my assignment through the portal	747	7.9%	6.8%	85.3%
Exchange of ideas through chats/instant messaging platforms	747	9.5%	10.7%	79.8%
Email communication with lecturers	747	22.6%	5.4%	85.4%
Use other online tools (Instant Messenger, Facebook, etc.) to contact lecturers about my studies	747	18.7%	14.2%	67.1%
Email communication with classmates	747	12.4%	9.9%	77.6%
Blogs/websites for sharing academic ideas with other students	747	11.6%	11.6%	76.7%
Blogs/websites for sharing academic ideas with other internet users	747	16.7%	12.4%	70.8%
Access to other learning materials for example, you tube videos	747	8%	7.8%	84.2%
Avails news prompts on the recent happenings in academia and technology	747	6.8%	13.4%	79.8%
Source of free online courses that have assisted me in a variety of disciplines	747	10.3%	11.9%	77.8%
Search online for available part time job opportunities for students	747	10.4%	11.9%	77.6%
Collection of a variety of online information good for my studies, then organizing it in files to be retrieved when I want	747	7.9%	11.2%	80.9%

3. METHODOLOGY

The prediction of student performance in an internetworked environment requires the collection of data regarding the factors highlighted in section 2: knowledge and competence levels of the learners in using internet technology, student effort in using the technology, the physical learning environment, attitude and behavior of the learners towards the technology, belief in the importance of the technology and the investment costs required to access internet technology. The collected data required a data mining technique that would assist in mining the data into useful information [27]. In educational data mining, prediction modelling has been used in the prediction of student performance. In order to build the predictive model, several tasks are involved. These are classification, regression analysis and categorization. The most popular approach to the prediction of student performance is classification.

The classification approach involves building structures of data from already existing data and these are used to make decisions from unseen cases [28]. The classification of the data is done in a two-step process. First, the model is built by analyzing the data list from the training data that has a set of attributes. The classification algorithm is then applied to the training data to create the model. Lastly, the final step of classification involves the use of test data to check the accuracy of the model. When the accuracy of the model is acceptable, then the model is used to classify the unknown data list [27].

There exists a number of algorithms under the classification task that have been used in the prediction of student performance. This paper focuses on the use of a decision tree.

3.1 Decision Tree
The algorithm implemented in this case is the decision tree algorithm. The decision tree has been used due to its simplicity and comprehensibility in uncovering both small and large data sets and predicting a value out of the data provided [29] [30] [31]. The decision tree has also been widely adopted in studies that have been used predict the performance of a student by considering a number of attributes, such as, final cumulative grade point average, psychometric factors, internal and external assessments. Consequently, decision tree models are easily understood because of their reasoning process and can be easily converted into IF-THEN statements [32].

A decision tree is a flow-chart-like tree structure, in which, internal nodes are denoted using rectangles and leaf nodes are denoted using ovals. All internal nodes have two or more child nodes. All internal nodes contain splits, which are used to test the value of an expression of the attributes. Arcs from an internal node to its children are labelled with distinct outcomes of the test. Each leaf node has a class label associated with it.

3.2 Construction of the Decision tree
The task of constructing a tree from the training set is known as the tree induction process or the tree building process. Most existing tree induction systems adopt a greedy (non-backtracking) top-down divide and conquer manner. Starting with an empty tree and the entire data set, a tree induction algorithm is applied on the training data (where each tuple is associated with a class label) until no more splits are possible [28]. The training set was constructed and the classification accuracy obtained from this dataset was 84.8485%.

3.3 Case study
The data was collected from Kenyan universities, 12 public universities and 8 private universities. Data collection involved the issuance of questionnaire surveys to students within these institutions. A total of 1,100 questionnaires were issued in the twenty universities. A total of 796 questionnaires were returned with 747 of them providing adequate information that could be used for data analysis. The research survey required the students to give their background information (category of university, education level and

gender), their knowledge on internet usage, their self-efficacy/capability with internet technology and their attitude towards the technology. The students were also required to survey other factors that enhanced their learning with internet technology, specifically, the physical environment and the general internet infrastructure and the influence by other parties (university peers, family and personality) to use the technology. The questionnaire also checked on the utility of internet technology, the factors that affect the usage of the technology, the relevance of the technology in the academic work of the student and the extent to which the technology enhanced the productivity of the student in the learning process.

The data was further refined into independent and dependent variables. The independent variables were classified as Perceived usefulness (PU), Perceived Ease of Use (PEOU), Task Technology Fit (TTF), Attitude (IAtt), Knowledge of Internet (KoI) and Investment (Inv). The dependent variable was classified as Performance (Perf).

Subsequently subjecting the independent variables (PU, PEOU, I Att, Inv, KoI and TTF) and the dependent variable (Perf) to a further grouping process assisted in combining the variables further. The grouping criteria was based on ensuring that less factors were used in the prediction model. Therefore, PU and IAtt were combined together to represent student effort (Effort). This was because, effort is dependent on behavioral intention and a set of actions performed. Behavioural intention depends on individuals' attitude and their belief in the usefulness of the technology (PU) [17].

The other set of factors that were combined together were TTF, PEOU and KoI. These factors affect the effectiveness of student effort in the utilization of internet technology in the learning process (Effe). For a student to use the technology effectively, they have to understand that the technology fits in the learning process, is relevant and useful in their studies (TTF and PEOU). In the same way, it's difficult to interact with a technology whose knowledge is not sufficient. Hence, the knowledge of the internet (KoI) will assist the student to successfully use their effort in the learning process. The final factor to be considered is investment (Inv). This factor is considered independent of any other factor since there are no other factors related to it.

The factors obtained after the grouping of the independent variable were student effort, investment and the effectiveness of student effort in the utilization of internet technology in learning. These were the factors used in the model construction.

3.4 Model Construction

In order to construct the student performance prediction decision tree, the data collected from the factors highlighted in 3.3 was consolidated together into one file, *Final8.arff*. This file was loaded on the WEKA explorer.

The WEKA workbench consists of a collection of visualization tools and algorithms for both analysis of data and prediction modelling. This is coupled up with graphical user interfaces for easy access to the prediction models and the analysis of data [33].

This workbench contains a number of panels that aid in the task of data mining. In this case, the *classify* panel is used because it helps in the application of the classification and regression algorithms on the dataset obtained. It also assists in the estimation of the accuracy of the prediction model obtained and helps to visualize any errors in the predictions given or in the model itself. There exists a number of decision tree algorithms that are used for classification on the WEKA workbench. The algorithm *classifier* that was used in this case was *J48*. This algorithm does not require discretization of the numeric attributes as compared to the other algorithms used on decision trees. In the *Test Options*, the *10-fold Cross-validation* is selected as the evaluation approach. This assists in getting an idea about the level of accuracy in the generated model.

3.5 Results obtained

The decision tree for the prediction of student performance after feature selection is shown on figure 1. The accuracy of the model is 84.6051%. This means that out of 747 instances of data collected, 632 instances are correctly classified. Therefore, the most important factor seen to influence the prediction of student performance in the utilization of internet technology is Student Effort. This factor forms the root of the decision tree.
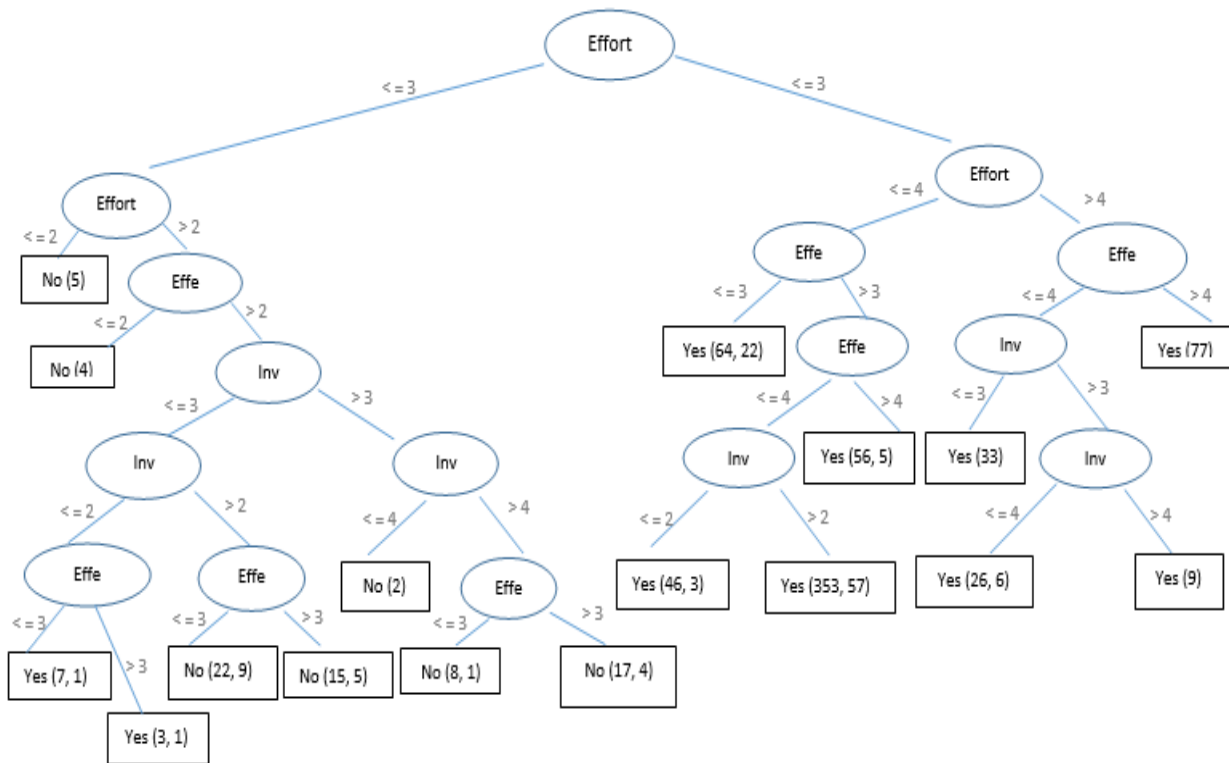
Figure 1. Decision tree of the factors that influence student performance in an internet mediated environment

The rules obtained from this tree are as follows:

1. *If Effort <=3 and Effort <=2 then Perf = no*

2. *If Effort <=3 and Effort > 2 else if Effe <= 2 then Perf = no*

3. *If Effort <=3 and Effe >2 else if Inv <=3 and Effe <=3 then Perf = no*

4. *If Effort <=3 and Effe >2 else if Inv <=3 and Effe >3 then Perf = yes*

5. *If Effort <=3 and Effe >2 else if Inv>2 and Effe<=3 then Perf = yes*

6. *If Effort <=3 and Effe>2 else if Inv >2 and Effe>3 then Perf = no*

7. *If Effort <=3 and Effe >2 else if Inv >3 or Inv <=4 and Effe <=3 then Perf = no*

8. *If Effort <=3 and Effe >2else if Inv >3 or Inv <=4 and Effe >3 then Perf = no*

9. *If Effort <=3 and Effe>2 else if Inv >3 or Inv >4 then Perf = no*

10. *If Effort >3 or Effort <=4 and Effe <=3 then Perf = yes*

11. *If Effort <=4 and Effe>3 else if Effe <=4 and Inv <=2 then Perf = yes*

12. *If Effort <=4 and Effe >3 else if Effe <=4 and Inv >2 then Perf =yes*

13. *If Effort >3 or Effert <=4 else if Effe >3 or Effe>4 then Perf = yes*

14. *If Effort >3 or Effort >4 else if Effe <=4 and Inv <=3 then Perf = yes*

15. *If Effort >3 or Effort >4 and Effe <=4 else if Inv >3 or Inv <=4 then Perf = yes*

16. *If Effort >3 or Effort >4 and Effe <=4 else if Inv >3 or Inv >4 then Perf = yes*

17. *If Effort >3 or Effort >4 and Effe > 4 then Perf = yes*

To elaborate further, for instance on rule 3:

If Effort <=3 and Effe >2 else if Inv <=3 and Effe <=3 then Perf = no

This means that, when student effort was low (<=3) then students' attitude and perceived usefulness of the technology were low (<=3). When the effectiveness of student effort in the utilization of internet technology was low (>2) then, this means that subjective norm to use the technology, the task-technology fit, the knowledge about the technology and the perceived ease of use of the technology were also low (>2). However, when the investment was low (<=3), this meant that availability and accessibility to the technology was also low. Hence, with a low value of the effectiveness of student effort in the utilization of internet technology in learning, the student perception about using the technology to improve their performance was low.

In the same way, to elaborate further, on another rule, for instance on rule 16:

If Effort >3 or Effort >4 and Effe <=4 else if Inv >3 or Inv >4 then Perf = yes

This means that, when student effort was high (>4), the effectiveness of student effort in the utilization of internet technology was high (<=4) and investment was high (>3), then student performance was perceived to be high. This means that, when student effort was high (>4), the student attitude and perceived usefulness of the technology were also high (>4). When the effectiveness of student effort in the

utilization of internet technology was high (<=4), the subjective norm, the task-technology fit, the knowledge about the technology and the perceived ease of use of the technology were also high (<=4). When investment was high (>4), this meant that availability and accessibility to the technology were at a high value (>4). Therefore, with a high value of investment in the technology, a high value in the effectiveness of student effort in the utilization of internet technology in learning and a high value of student effort, the student perception about using the technology to improve their performance was high.

Another way that is used to summarize the prediction results [34], is based on the confusion matrix shown in table 7. This matrix gives the number of the correct and incorrect classifications that were obtained from the *class Perf.*

Table 7. Confusion matrix

= = = Confusion Matrix = = =		
a	b	← classified as
61	93	a = no
22	571	b = yes

In table 7, the total number of actual **no's** in the dataset considered is the sum of the values on the **'a'** column (61+22=83). The total number of actual **yes's** in the dataset is the sum of the values on the **'b'** column (93+571=664). The total number of correct values are organized in a diagonal line from the top-left to the bottom-right on the matrix (61+571= 632). The total number of incorrect values are organized as the remaining values on the matrix (93+22=115). Hence, this means that more errors were made in predicting the **yes's** than the **no's**.

4. CONCLUSION
The student performance prediction model has assisted in establishing the major factors required in the prediction of student performance within internet mediated environments. As a result, the different factors have been used to generate the decision tree which predicts the student performance perceptions based on the values of the different factors. From the models accuracy level of 84.6051%, it is clear that the model can be used to successfully predict the performance perceptions of students who use internet technology in their learning process.

5. ACKNOWLEDGEMENT

6. REFERENCES
[1] Hughes, G., & Dobbins, C. (2015). The utilization of Data Analysis Techniques in Prediction Student Performance in Massive Open Online Courses (MOOCs). *Research and Practice in Technology Enhanced Learning, 10*(10).

[2] Chen, M., Mao, S., & Liu, Y. (2014). Big data : a survey. *Mobile Networks and Applications, 19*(2), 171-209.

[3] Ramesh, V., Parkavi, P., & Ramar, K. (2013). Predicting student performance: a statistical and data mining approach. *International Journal of Computer Applications, 63*(8), 35-39.

[4] Brews, P. J., & Tucci, C. L. (2015). *Exploring the performance effects of Internetworking.* CSI-REPORT-200 5-002.

[5] Mills, K. (2016). Review Possible Effects of Internet Use on Cognitive Development in Adolescence. *Media and Communication, 4*(3), 4-12. doi:10.17645/mac.v4i3.516

[6] Nwokedi, V. C. (2010). Impact of Internet Use on Teaching and Research Activities of the Academic Staff of Faculty of Medical Sciences: A case study of University of Jos. *Gateway Library Journal, 10*(1), 13-22.

[7] Carbonaro, W. (2005). Tracking, Students' Effort, and Academic Achievement. *Sociology of Education*, 27–49.

[8] Skinner, E., & Pitzer, J. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21-44). New York, NY: Springer.

[9] Reschly, A., & Christenson, S. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3-19). New York,NY: Springer.

[10] Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387. doi:10.1037/a0026838

[11] Lizzio, A., Wilson, K., & Simons, R. (2002). University Students' Perceptions of the Learning Environment and Academic Outcomes: implications for theory and practice. *Studies in Higher Education, 27*(1).

[12] Murugan, A., & Rajoo, L. (2013). Students' Perceptions of Mathematics Classroom Environment & Mathematics Achievement: A Study in Sipitang, Sabah, Malaysia. *Proceeding of The International Conference on Social Science Research, Icssr.* Penang, Malaysia.

[13] Ashby, J., Sadera, W., & S, M. (2011). Comparing Student Success Between Developmental Math Courses Offered Online, Blended, And Face-To-Face. *Journal of Interactive Online Learning, 10*(3).

[14] Frenzel, A., Pekrun, R., & Goetz, T. (2007). Perceived Learning Environment and Students' Emotional Experiences: A Multilevel Analysis of Mathematics Classrooms. *Learning and Instruction, 17*.

[15] Shamaki, T. A. (2015). Influence of Learning Environment on Students' Academic Achievement in Mathematics: A Case Study of Some Selected Secondary Schools in Yobe State – Nigeria . *Journal of Education and Practice, 6*(34).

[16] Fuchs, W. (2010). Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School. *CESifo*.

[17] Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood

Cliffs, NJ: Prentice Hall Inc.

[18] Ajzen, I. (1991). The theory of planned behavior. *Organization of Behaviour in the Human Decision Process, 50*(2), 179–211.

[19] Trafimow, D., & Lench, H. (2015). The sufficiency assumption of the reasoned approach to action. *Cogent Psychology, 2*(1), x.

[20] Mathieson, K. (2000). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. *Information System Research, 2*(3), 173-191.

[21] Bidin, Z., Mohd-Shamsudin, F., Hashim, M. F., & Sharif, Z. (2011). Factors Influencing Students' Intention to Use Internet for Academic Purposes. *SSRN Electronic Journal*. doi:10.2139/ssrn.1867902

[22] Mitra, A., & Steffensmeier, T. (2010). Changes in student attitudes and student computer use in a computer-enriched environment. *Journal of Research on Technology in Education, 32*(3), 417- 433.

[23] Hong, K.-S., Ridzuan, A. A., & Kuek, M.-K. (2017). Students' attitudes toward the use of the Internet for learning: A study at a university in Malaysia. *Journal of Educational Technology & Society*, 45-49.

[24] Croteau, D., & Hoynes, M. (2013). *Media Society: Industries, Images and Audiences* (10th ed.). Pine Forge Press, Thousand Oaks.

[25] Hussain, A., Mkpojiogu, E., & Yusof, M. M. (2016). Perceived usefulness, perceived ease of use, and perceived enjoyment as drivers for the user acceptance of interactive mobile maps. *AIP Conference Proceedings.* American Institute of Physics. Retrieved from https://doi.org/10.1063/1.4960891 View Table of Contents: http://aip.scitation.org/toc/apc/1761/1

[26] Huesemann, M., & Huesemann, J. (2011). *Technofix: Why Technology Won't Save Us or the Environment, "The Myth of Value-Neutrality".* Gabriola Island, British Columbia, Canada: New Society Publishers.

[27] Kabra, R., & Bichkar, R. (2011). Performance Prediction of Engineering Students using Decision Trees. *International Journal of Computer Applications, 36*(11).

[28] Han, J., & Kamber, M. (2011). *Data Mining Concepts and Techniques* (2nd ed.). Morgan Kaufmann, SanFrancisco.

[29] Osmanbegovi´c, E., & Sulji´c, M. (2015). Data mining approach for predicting student performance. *Economic Review, 10*(1).

[30] Natek, S., & Zwilling, M. (2014). Student data mining solution–knowledge management system related to higher education institutions. *Expert systems with applications, 41*(14), 6400–6407.

[31] Shahiria, A. M., Husaina, W., & Rashida, N. A. (2015). A Review on Predicting Student's Performance using Data Mining Techniques. *The Third Information Systems International Conference.*

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

[33] Witten, I., & Frank, E. (2011). *Data Mining: Practical Machine Learing Tools and Techniques with Java Implementations.* San Diego, CA: Morgan Kaufmann.

[34] Brownlee, J. (2016). *Weka Machine Learning*. Retrieved from How to Perform Feature Selection With Machine Learning Data in Weka.