IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

Correlation and Linear Regression Analysis (CALRA): A Predictive Decision Support System for LEAE

Domingo V. Origines Jr[1], Dr. Meliza P. Alo[2]

Information Technology Department, SPAMAST,
Matti, Digos, Davao Del Sur, Philippines[1]

Professor, Graduate School, Southern Philippines, Agri-business, Marine, and Aquatic
School of Technology (SPAMAST), Matti, Digos, Davao Del Sur, Philippines[2]

Abstract: This study introduce a Decision Support System called Correlation and Linear Regression Analysis – Decision Support System (CALRA-DSS) System. The CALRA-DSS will use Data mining technique in the process of discovering knowledge which in turn can be used to predict future results. CALRA-DSS predicts Students chances whether passed or Failed in the LEAE Licensure Examination. The integration of Data Mining Technique using Pearson-Product Moment Correlation that is used to determine the degree to which two variables are related and Regression Analysis that is used to examine the relationship between and one dependent and one independent variable. The data to be tested by this CALRA-DSS were the April 2016 Agricultural Engineering graduates of ICET, SPAMAST- Digos Campus who participated in the LEAE Mock Board Examination last and took the August 2016 LEAE. The academic records of these graduates were taken from the SPAMAST Electronic Students Information System (eSMS) Digos Campus while the Mock Board data Result was taken from the SPAMAST LEAE Reviewer Committee. It is concluded that with the use of this tool, the ICET Department can implement an intervention program timely before Student would intend to take the LEAE. Based on the Outcome, the results obtained from the Correlation and Regression Analysis and the attributes obtained from eSMS, the identified Academic Predictors has a strong correlation to Mock Board Examination. In general outcome of the study can give a hint to the Students as to which subjects can be considered as predictor variables for their licensure exam scores and hence become their focus of study/review while still studying.

Keywords: Correlation, Data Mining, Regression, LEAE, Academic Performance, Prediction.

I. INTRODUCTION

Decision support systems (DSS) are defined as interactive application systems intended to help decision makers utilize data and models in order to identify problems, solve problems and make decisions. The mission of decision support systems is to improve effectiveness, rather than the efficiency of decisions [21]. This study introduce a Decision Support System called Correlation and Linear Regression Analysis–Decision Support System (CALRA-DSS) System, a system tool that can process and discover a knowledge which in turn can be used to predict future results. Using Academic Predictors, the CALRA-DSS can predict Students chances whether passed or Failed in the Licensure Examination. The integration of Data Mining Technique to the system using Pearson-Product Moment Correlation that is used to determine the degree to which two variables are related and Regression Analysis that is used to examine and predict the relationship between one dependent and one independent variable. Basically, regression takes a numerical dataset and develops a mathematical formula that fits the data. After performing an analysis, the regression statistics used to predict the dependent variable when the independent variable is known [6]. The integration allows the use of more than one input variable and allows for the fitting of more complex models basis for Decision Support System that strongly predicts the relationship between academic variables and result in the Mock Board Examination[9].

In an Academic Institution using Decision Support System, there are several ways of defining the quality of higher education institutions (HEI) in the Philippines. One tangible measure commonly used in the country is the performance of an HEI's graduates in state licensure examinations [7]. There have been several attempts to discover models in predicting the performance in licensure examination but most studies recommend for an extensive study covering other independent variables and other approaches [8].

Garciano found out that the Academic Performance such as General Education Subjects, Agricultural Engineering (AE) Major Subjects, and 80% Score Performance on the Mock Board Examination has a strong correlation in achieving a higher level of passing rate in the Licensure Examination for Agricultural Engineering (LEAE) to the Graduates of BSAE in SPAMAST [1]. It was also significantly revealed by Arce, S. E. and Belen, J. L. that the relationship of In-House review to LET performance using descriptive – correlational method [5] and found out that pre-board and LET results are significantly correlated.

In SPAMAST, the Licensure Exam for Agricultural Engineering is under the Institute of Computing Engineering and Technology (ICET). One of the pressing concerns of the ICET of the Southern Philippines Agri-business and Aquatic and School of Technology (SPAMAST) is the batting or passing percentage of its graduates in the Licensure Examination for Agricultural Engineering (LEAE)[1]. Based on the licensure examination result from Philippine Regulation Commission (PRC), SPAMAST has experienced low performance in Agricultural Engineering Board Program offered from 2006 up to Present. In order to improve and increase the Passing Percentage for the LEAE in SPAMAST, A User-friendly tools in predicting the Students chances whether passed or Failed in the Licensure Examination must be utilized by the ICET Department so that the ICET could Implement timely an intervention program before the Students would take the Licensure Examination.

II. REVIEW OF RELATED LITERATURE

Several reform agenda and programs have been initiated to set directions for State Universities and Colleges (SUCs) for them to catch up with top universities and colleges in other ASEAN countries. Among these are the Philippine Development Plan 2011-2016; Public Higher Education Reform Agenda issued in 2012; the criteria in evaluating the performance of SUC under Executive Order (EO) No. 80, s. 2012; and CHED Memorandum Order (CMO) No. 46, s. 2012 on typology and outcomes-based quality assurance. The results of these initiatives should now be quantifiable in terms of the performance of issues in the areas of quality and relevance of instruction, research capability and output, services to the community, and management of resources [7].

The 113 SUCs are at various stages of development. Many of them have outgrown their levels both quantitatively and qualitatively. Thus, there is a need to conduct another leveling exercise to determine the current status of SUCs and to level the playing field with private higher education institutions (HEIs). I . The Fiscal Year (FY) 2016 leveling instrument for SUCs has been prepared jointly by the CHED and the DBM, in coordination with the PASUC. One of the Indicators for SUC Leveling is the Performance of Licensure Examination on Board Programs being offered by the SUCs [7].

There are several ways of defining the quality of higher education institutions (HEI) in the Philippines. One tangible measure commonly used in the country is the performance of an HEI's graduates in state licensure examinations. Through these examinations, the skills, and competencies which are said to be the likely outputs of a quality-assured HEI could be measured to some extent basis for SUC's leveling [2].

BS in Agricultural Engineering is a five-year program that concentrates on the technical aspect of Agriculture. Its curriculum revolves around the application of engineering principles in the production, processing, handling and storage of food, fiber, and materials. It also covers areas such as irrigation and drainage of agricultural land, soil erosion control, the planning of agricultural buildings and structures, post-harvest technology and agricultural waste management. It has a board examination. The subjects in the BSAE curriculum are divided into the following categories: General Education courses, Fundamental Agricultural courses, Basic Engineering courses and Professional courses[2][7].

To become a Licensed Agricultural Engineer in the Philippines, a graduate of BS in Agricultural Engineering needs to pass the Agricultural Engineer Licensure Examination. The examination is conducted by the Board of Agricultural Engineering under the supervision of the Professional Regulations Commission (PRC). The exam is conducted once a year [2][7].

The PRC Licensure examination for Agricultural Engineers (LEAE) covers the following subjects. Thirty Percent (30%) for Subject 1 covers topics in Agricultural mechanization, Power Machinery, and Equipment. Another thirty percent (30%) for Subject 2 which covers Soil and Water Conservation, Irrigation and Drainage and another forty percent (40%) for subject 3 which covers concerning Rural Electrification, Agricultural Processing and Agricultural Statistics [2].

There have been several attempts to discover models in predicting the performance in licensure examination but most studies recommend for an extensive study covering other independent variables and other approaches [8]. A Data Mining Technique is a way to discover new meaning in data, perform data processing using sophisticated data search capabilities and statistical algorithms, which can be utilized in any organization or system that needs to determine the patterns or relationship implicit in a large data warehouse for better strategies to best reach them[6]. Data Mining was originally developed to act as expert systems to solve problems and did not require assumptions to be made about data[8].

Hand et al, cited that in a real world, predicting the performance of the students is a challenging task. The primary goals of Data Mining in practice tend to be Prediction and Description. Predicting performance involves variables like Math, Programming language, Lab Marks etc. in the student database to predict the unknown or future values of interest. Educational Data Mining uses many techniques such as Correlation, Regression, Decision Trees, MultiLayer perception

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

NaïveBayes and many others. Using these methods many kinds of knowledge can be discovered[14].
Lee and Park presented the Customized Sampling Decision Support System (CSDSS) which uses data mining [12]. CSDSS is a web-based system that enables the user to select a process sampling method that is most suitable according to his needs at purchasing semiconductor products. The system enables the autonomous generation of the available customized sampling methods and provides the performance information for those methods. CSDSS uses clustering data mining method within the generation of sampling methods.
Zupan et al., also tried to bridge the data mining and decision support in their study. Specifically, they proposed a mechanism for communicating data mining models. The predictive model was developed separately within a data mining tool which is called Orange, then the model is encoded in XML [18] hence, models are hard coded and fixed. Baking and Quimbao introduced the same concept by proposing a knowledge-driven educational decision support system for education with a semester credit system by taking advantage of educational data mining [5]. This is, however, different from this paper's objective though it was considered to be in the educational context.
Meanwhile, Mock Examinations defined as one of the predictors that influenced the performance of Education Graduates [9]. Bachelor of Science in Basic Education Graduates of the University of the Cordilleras have been studied through undergoing mock examination in the L.E.T. Imitating the actual examination, College of Education Faculty organized the comprehensive examination and it has found out that respondents gained positive impact of the said examination in their performance. It has been recommended to continually conduct comprehensive or mock examinations to education graduates before taking Licensure Examination for Teachers to increase their probability in passing the board examination [10].
Arce, S. E. and Belen, J. L. undertook a study that revealed the relationship of In-House review to LET performance using descriptive – correlational method [5] and found out that pre-board and LET results are significantly correlated.
On the other hand, the profile of respondents, the level of motivation, gender, race and career preferences are personal factors that serve as the strong determinants of Board Examination Performance [19], and (17). Mental and Psychological enhancement of education graduate through a high level of motivation, perseverance; determination and influence by the institution to pass the board examination have been shaped as one of the strong factors affecting Licensure Examination for Teachers. Based on the study conducted by [19] in University of Northern Philippines, female respondents achieved a higher level of passing rate in board examination and most of them came from Music, Arts, Physical Education and Health (MAPEH) major and examinees' Grade Point average ranged from 2.00-2.25. Likewise, institutions that achieved high level of passing rate serves as the influential factor in the personalities of examinees to pass the board examination [19][17].
Previous studies have indicated that by and large, performance of Education Graduates in Licensure Examination for Teachers is shaped both by personal and educational factors. As an evidence, a study conducted by Filipino researchers have been found that academic performance in terms of Specialization (Computer Education, English, Filipino, Mathematics, Science and Social Studies), General Education (Social Sciences, Mathematics, Science, Filipino and English) and Professional Education serves as the strong predictor of Board Examination Performance [19], [4],and [10]. Also, Admission Test, Degree Course, English Proficiency and Institutional Passing rate of Education Graduates are the factors affecting with the effectiveness of education respondents in Licensure Examination for Teachers [19]. Teaching factors and implemented policies of In-house review conducted by universities and review center are the factors that largely affect their board examination performance [20] .
All the above-cited works gave credit to the importance of integrating decision support system and data mining. So far, it was undertaken that focus on student records as a Predictors basis for DSS will lead to a good Passing Performance rate for future Student Licensure Examination.

III. THEORETICAL FRAMEWORK

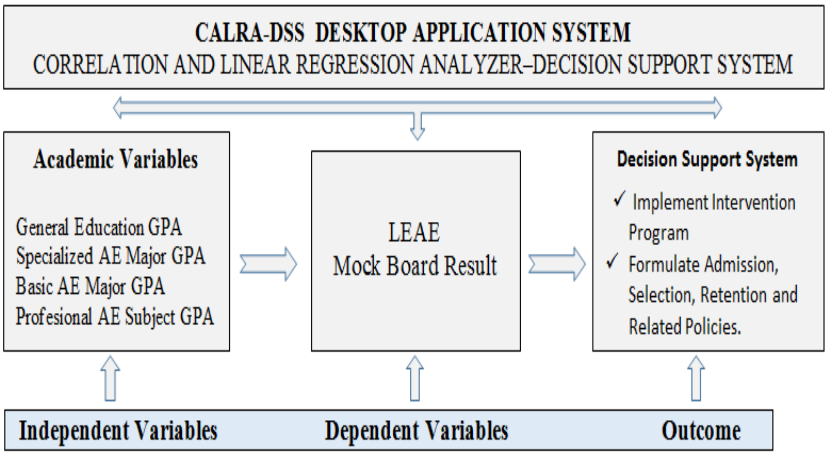


Figure 1. Conceptual Framework

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

Figure 1 shows the Conceptual Framework of the Study. The GPA for General Education, Specialized AE Major, Professional Courses and Basic AE Subject Courses are the Independent Variables while Mock Board Examination Result is the Dependent Variables. The Independent and Dependent Variables would be the basis for the Decision Support System as Outcome variable. The Independent, Dependent and Outcome Variables are the components of CALRA-DSS Desktop Application System.

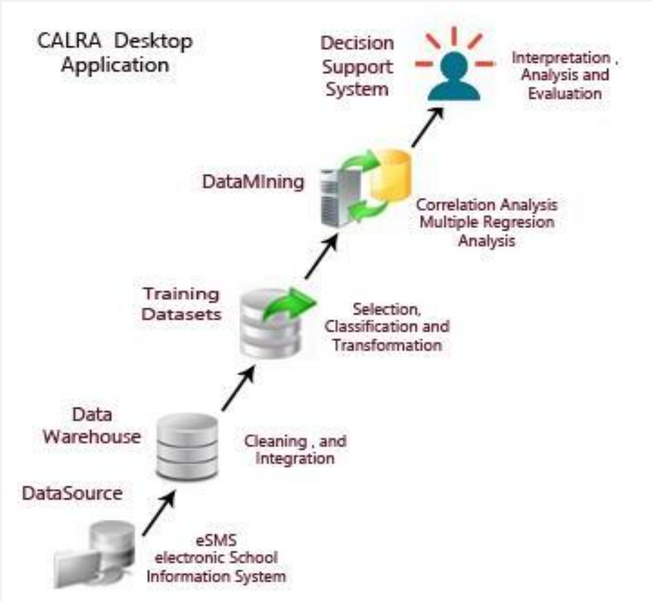


Figure 2. Data Mining Framework

Figure 2 shows the Data Mining Framework. It follows from **Data Sources** which will perform downloading of Data from eSMS, **Data Warehouse** which will perform Cleaning and Integration, **Training Datasets** which will do the Selection, Classification and Transformation, **Training Datasets** which will perform Correlation and Regression Analysis, and **Decision Support System** of which the System User will do Interpretation, Analysis and Evaluation.

IV. OPERATIONAL FRAMEWORK

Methodology

The data to be tested by this CALRA-DSS were the April 2016 Agricultural Engineering graduates of ICET, SPAMAST- Digos Campus who participated in the LEAE Mock Board Examination last July 2016 and took the August 2016 LEAE. The academic records of these graduates were taken from the SPAMAST Electronic Students Information System (eSMS) Digos Campus while the Mock Board data Result was taken from the SPAMAST LEAE Reviewer Committee. These data that were stored in different tables were uploaded and cleaned by removing duplicate records. Records that contain empty values were likewise deleted. The integration was done to the different tables into one data warehouse, the data of which were transformed to create meaningful groups within the attributes to match that of the objectives of the study. The Datasets uploaded from the eSMS to CALRA-DSS were categories according to General Education Subjects (GPA_GenEd), Basic Agricultural Engineering Subjects (GPA_Basic), Professional Engineering Subjects (GPA_Prof) and Specialized Major Core Subjects (GPA_Special).

Specifically, the CALRA-DSS system has the following functionalities:

I. The system had an option to Connect to the eSMS Software and Download the Student Grades Data. It has also an option to upload Data in the form of text files or excel files.
II. The downloaded Data will categorized according to the following classification Attributes: GPA_GenEd, GPA_Basic, GPA_Major, and GPA_Special.
III. The System User Administrator can Create a Mock Board Questionnaire and the functionality to Generate and Compute a Summary Results per individual Student Takers.
IV. The CALRA-DSS can generate Correlation Analysis and Regression Analysis based on the Record set downloaded and Classified per Individual Recordset.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

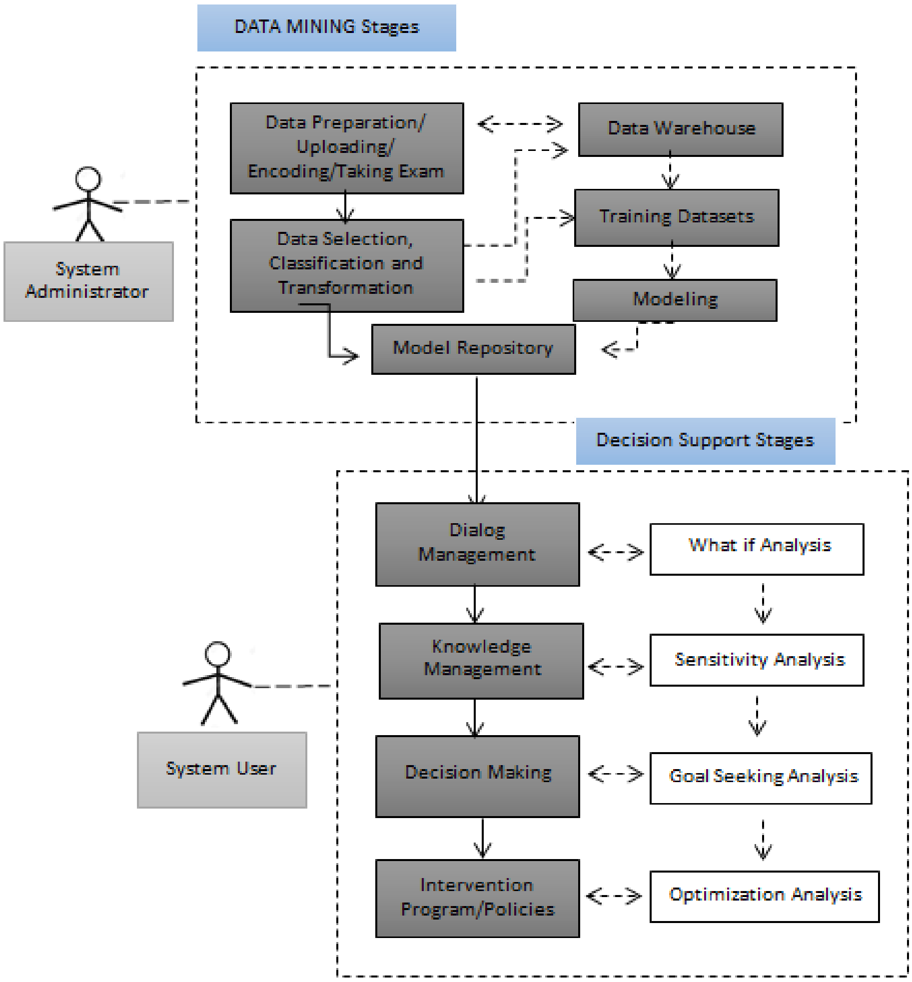


Figure 3. Process Model for Data Mining and Decision Support System

Figure 3 Show the process model for Data Mining and Decision Support System. It has two Stages the Data Mining Stages consist of Data Preparation and Uploading for Data Warehousing on Training Datasets that will produce Modeling into the Modeling Repository basis for Decision Support System. The Decision Support Stage consist of Dialogue Management of which what if analysis should be done, Knowledge Management which need sensitivity analysis, decision making that needs goal seeking analysis and Intervention on Program/Policies that needs optimization analysis.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

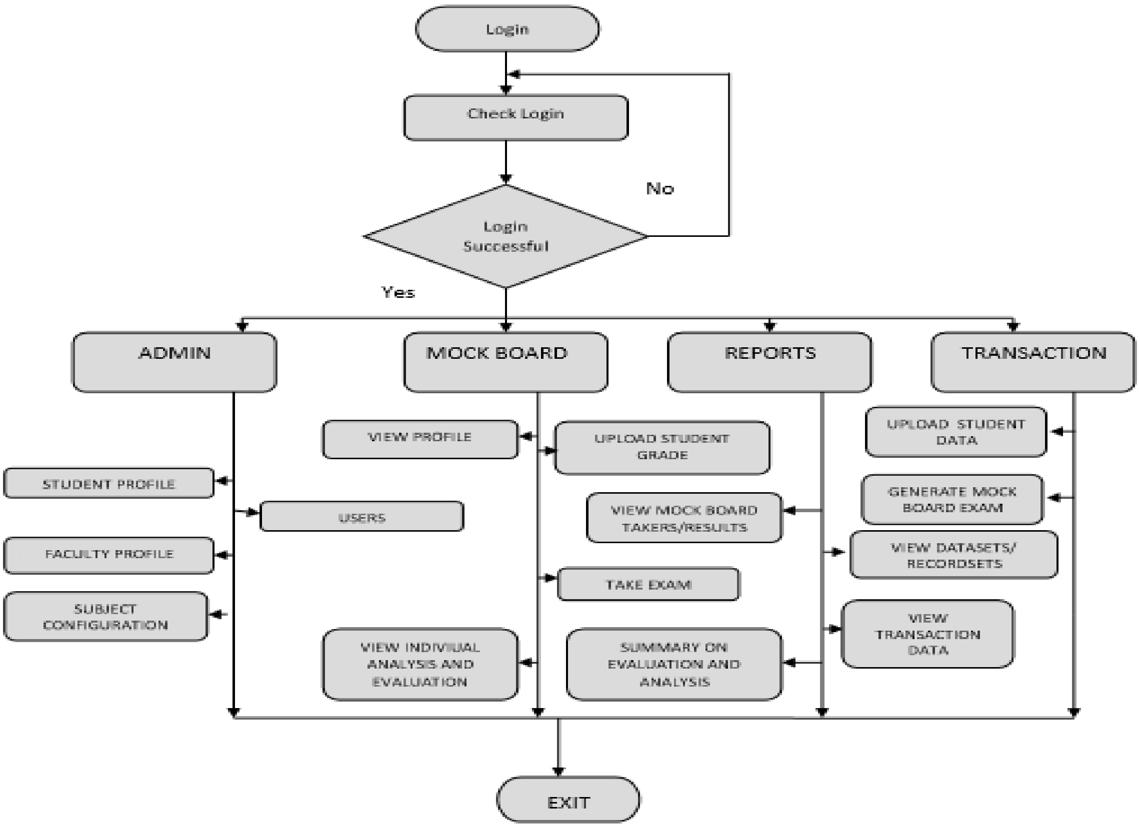


Figure 4. Data Flow Diagram

The Data Flow Diagram of CALRA-DSS. The System User needs to log-in, after validated by the system the Main Interface will appear. The Main Interface has Admin Menu, Mock Board Menu, Reports Menu and Transaction Menu as show in figure 4.

.

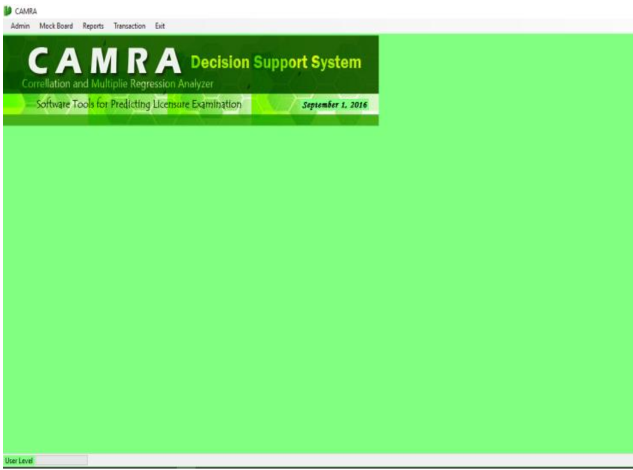


Figure 5. Main Interface for CALRA-DSS

Figure 5 shows the Main Interface for CALRA-DSS. It has **Admin Menu** that has sub menu for Student Profile, Subject configuration, and User Data. The **Mock Board Menu** has sub-menu for Take Exam, User Profile, and Individual Summary Result. The **Report Menu** has sub-menu for View Mock Reports, Summary, Summary on Board Questionnaire, and Individual Summary Results. The **Transaction Menu** has Upload Data and Generate Mock Board sub-menu.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

V. TESTING AND SIMULATION

Correlation Analysis using Pearson Correlation Coefficient

A Correlation Analysis using Pearson-Product Moment Correlation (r) was integrated to the CALRA-DSS to determine the relationship between identified Academic Variables and Result in the Mock Board Examination for Agricultural Engineering Students.

In order to determine how strong the relationship is between two variables, a formula must be followed to produce what is referred to as the **coefficient value**. The coefficient value can range between -1.00 and 1.00. If the coefficient value is in the negative range, then that means the relationship between the variables is **negatively correlated**, or as one value increases, the other decreases. If the value is in the positive range, then that means the relationship between the variables is **positively correlated**, or both values increase or decrease together. The formula for conducting the Pearson correlation coefficient value [3].

The Datasets that is obtained from eSMS Software and generally converted into Percentage Equivalent as shown in Table 1 , so that further calculations can be made on the data and it becomes easy to handle.

TABLE 1. NUMERICAL PERCENTAGE RANGE OF PREDICTORS

GPA Range per Predictors	Percentage Equivalence	Qualitative Rating
1.0 -1.24	99 - 100	Excellent
1.25-1.4	96 - 98	Outstanding
1.5-1.74	93 - 95	Good Work
1.75.1.4	90 - 92	Satisfactory Work
2.0-1.74	87 - 89	Average
2.25-1.9	84 - 86	Moderately Average
2.5-2.24	78 - 83	Moderately Low Average
2.75-2.4	74 - 77	Low Average
3.0-2.74	70 - 73	Passing

Table 1 show the GPA ranges Predictors, Percentage Equivalence and the quantitative ratings for the Academic Predictors. The Equivalence will be used in converting Decimal Point Grades into Percentage Equivalence.

TABLE 2. STUDENT GRADE DATA AS DATA SOURCE DOWNLOADED FROM ESMS SOFTWARE

StudID	**Course**	**Course Code**	**Semester**	**School Year**	**Course Description**	**Units**	**FGrade**	**Category**
2012-05001	BSAE	Eng 1	1st Semester	2012-2013	Study and Thinking Skills	3	1.75	GPA_GenEd
2012-05001	BSAE	Physics 1a	1st Semester	2012-2013	Mechanics and Heat	3	2.25	GPA_Basic
2012-05001	BSAE	Chem 2a	1st Semester	2012-2013	Organic Chemistry	3	2.25	GPA_GEnEd
2012-05001	BSAE	Eng 2	1st Semester	2012-2013	Writing in the Discipline	3	2.25	GPA_GEnEd
2012-05001	BSAE	Fil 2	1st Semester	2012-2013	Pagbasa at Pagsulat sa Iba't-ibang Disiplina	3	1.5	GPA_GEnEd
2012-05001	BSAE	Math 3	1st Semester	2012-2013	Analytic Geometry and Calculus I	3	2.25	GPA_Basic
2012-05001	BSAE	PE 2	1st Semester	2012-2013	Rhythmic Activities	2	1.75	GPA_GenEd
2012-05001	BSAE	NSTP 2	1st Semester	2012-2013	LTS/CWTS/ROTC 2	3	1.5	GPA_GenEd
2012-05001	BSAE	Hist 1	2nd Semester	2013-2014	Philippine History	3	1.5	GPA_GenEd
2012-05001	BSAE	Hum 1	2nd Semester	2013-2014	Introduction to Humanities	3	2.0	GPA_GenEd

Table 2 shows the downloaded Student Grades from eSMS Software to be used as Data Source for the Academic Predictors.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

TABLE 2a. Average Gpa For Each Category From Data Source

STUD_ID	GPA_ GenEd	GPA_ Basic	GPA_ Prof	GPA_ Special	GPA_ OverAll
2012-05001	95	92	90	94	93
2012-05002	86	85	85	87	86
2012-05003	90	90	88	85	88
2012-05004	80	82	81	79	81
2012-05005	83	90	89	86	87
2012-05006	87	88	84	92	88
2012-05007	84	82	83	86	84
2012-05008	71	73	75	77	74
2012-05009	70	72	77	75	74
2012-05010	77	79	76	79	78

Table 2a shows the Average GPA for each Category downloaded from Data Source. It is now transformed by the CALRA-DSS System into Percentage equivalent per Category.

Table 2b. Result from Student Mock Board Examination per Category

STUD_ID	GenEd_ Result	Basic_ Result	Prof_ Result	Special_ Result	Avg_ Mock Board Result
2012-05001	90	90	91	97	92
2012-05002	87	84	85	88	86
2012-05003	89	86	88	87	88
2012-05004	67	66	80	80	73
2012-05005	88	88	89	88	88
2012-05006	84	86	84	87	85
2012-05007	88	81	83	86	85
2012-05008	45	56	76	67	61
2012-05009	78	71	72	76	74
2012-05010	56	55	59	60	58

Source: Result Mock Board Examination last May 2016 from SPAMAST Review Center.

Table 2b shows the Average GPA Mock Board for each Category gathered from SPAMAST Review Center last May 2016.

Table 3. Consolidated GPA_Overall and Average Mock board Result

STUD_ID	GPA_OverAll	AVG_MockBoard Result
2012-05001	93	92
2012-05002	86	86
2012-05003	88	88
2012-05004	81	73
2012-05005	87	88
2012-05006	88	85
2012-05007	84	85
2012-05008	74	61
2012-05009	74	74
2012-05010	78	58

Table 3 shows the consolidated result from GPA_Overall as Independent Variables and AVG Mock Board Result as Dependent Variables.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

TESTING FOR CORRELATION ANALYSIS

Table 4 shows the Predictors data for two variables, labeling the variables (*x*) and (*y*), and add three more columns labeled (*xy*), (*x*^2), and (*y*^2).

Table 4. Combined Datasets for GPA_OverAll and Mock Board Result as x and y Variables

STUD_ID	GPA_ OverAll (y)	AVG_ Mock Board Result (x)	xy	x^2	y^2
2012-05001	93	92	8,533	8,464	8,603
2012-05002	86	86	7,375	7,396	7,353
2012-05003	88	88	7,722	7,656	7,788
2012-05004	81	73	5,897	5,366	6,480
2012-05005	87	88	7,678	7,788	7,569
2012-05006	88	85	7,481	7,268	7,700
2012-05007	84	85	7,077	7,140	7,014
2012-05008	74	61	4,514	3,721	5,476
2012-05009	74	74	5,457	5,513	5,402
2012-05010	78	58	4,471	3,306	6,045
∑	831	790	66,203	63,618	69,430

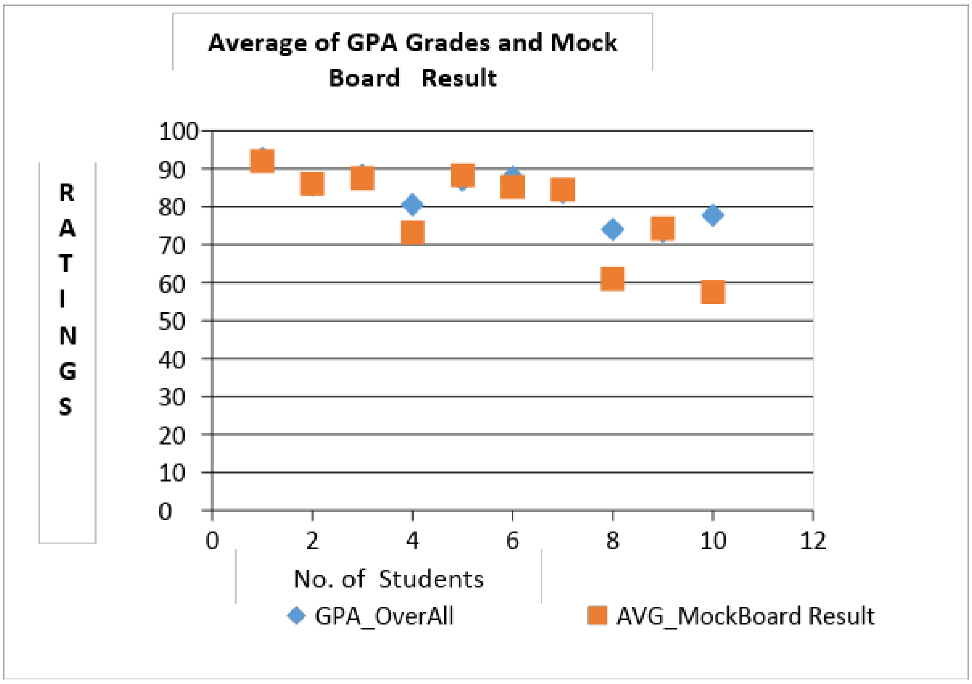


Figure 6. Scatter Plot for x and y Variables

The correlation coefficient formula is as follows:

(r) =[nΣxy – (Σx)(Σy) / Sqrt([nΣx2 – (Σx)2][nΣy2 – (Σy)2])]

r: The correlation coefficient is denoted by the letter r.

n: Number of values. If we had ten(10) people we were calculating the correlation coefficient for, the value of n would be 10.

x: This is the first data variable.

y: This is the second data variable.

Σ: The Sigma symbol (Greek) tells us to calculate the "sum of" whatever is tagged next to it.

DOI 10.17148/IJARCCE.2019.8401

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

Pearson's Correlation Coefficient is a linear correlation coefficient that returns a value of between -1 and +1. A -1 means there is a strong negative correlation and +1 means that there is a strong positive correlation. A 0 means that there is no correlation (this is also called **zero order correlation**).

Table 5. Range of P interpretation

Value of ρ	Strength of relationship
-1.0 to -0.5 or 1.0 to 0.5	Strong
-0.5 to -0.3 or 0.3 to 0.5	Moderate
-0.3 to -0.1 or 0.1 to 0.3	Weak
-0.1 to 0.1	None or very weak

Formula for Correlation Coefficient and Result:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

$$= \frac{10(66{,}203) - (790)(831)}{\sqrt{[10(63{,}618) - (790)^2][10(69{,}430) - (831)^2]}}$$

$$= 0.856$$

where:

r = Sample correlation coefficient
n = Sample size
x = Value of the independent variable
y = Value of the dependent variable

Table 6. Using Microsoft Excel (Excel Correlation Output)

	GPA_OverAll	*AVG_MockBoard Result*
GPA_OverAll	1	
AVG_MockBoard Result	0.858549994	1

The formula above shows that after plugging in all the correct values, the result is the coefficient value! If the value is a negative number, then there is a negative correlation of relationship strength, and if the value is a positive number, then there is a positive correlation of relationship strength as shown in Table 6.

Figure 6 shows the Scatter Plot for GPA_OverAll and AVG MockBoard Result Examination. It also shows from Table 6, the Computed Data using the Pearson correlation coefficient value from the formula below. Result shows from the Formula and Table 5 as result done in Excel Correlation Output, that there is a relatively strong positive linear association between x and y variables, meaning the null hypotheses is rejected.

To do the other test with regards to the null hypothesis, H0, that there is no correlation in the population against the alternative hypothesis, H1, that there is correlation; our data will indicate which of these opposing hypotheses is most likely to be true. We can thus express this test as:

Significance Test for Correlation

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

- Hypotheses
  Ho_0: ρ = 0 (no correlation)
  HA: ρ ≠ 0 (correlation exists)
- Test statistic
  – (with n – 2 degrees of freedom)

Hypotheses: *There is no significant relationship between identified academic Variables and Performance in the Mock Board examination*

Result:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{.856}{\sqrt{\frac{1-.856^2}{10-2}}} = 4.887$$

Conclusion: There **is evidence** of a linear relationship at the 5% level of significance. The Decision is Reject $H_{0.}$

DOI 10.17148/IJARCCE.2019.8401

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

A. REGRESSION ANALYSIS

The Regression Analysis predicts the value of a dependent variable based on the value of at least one independent variable. In this Regression Analysis the Data from Table 4 was used as x and y Variables. The formula for Regression Analysis is shown in below.

Estimated (or predicted) y value

Estimate of the regression intercept

Estimate of the regression slope

Independent variable

$$\boxed{\hat{y}_i = b_0 + b_1 x}$$

$$b_1 = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2} \qquad b_0 = \bar{y} - b_1\bar{x}$$

Figure 7. Regression Analysis Formula

b_0 is the estimated average value of Y when the value of X is zero (if x = 0 is in the range of observed x values)

Using the Microsoft Excel for testing, the coefficients b_0 is the regression intercept while b_1 is the slope. The Table 7 shows the Summary Output on Regression Analysis using Data from Table 4 using MockBoard Result as x Variable and GPA_OverAll as y Variables.

Regression Statistics	
Multiple R	0.858549994
R Square	0.737108092
Adjusted R Square	0.704246604
Standard Error	3.507026318
Observations	10

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	275.88113	275.881	22.430758	0.001471
Residual	8	98.393869	12.2992		
Total	9	374.275			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	46.546575	7.797295	5.96958	0.0003354
AVG_MockBoard Result	0.46299462	0.0977584	4.73611	0.001473
	Lower 95%	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
	28.56598	64.52717	28.566	64.527
	0.2375634	0.6884258	0.23756	0.6884

Given the Result Summary as shown in Table 7, the Estimated Regression Formula derived from the Regression Summary output will have the values below:

AVG_MockBoard Result = 46.54657482 + 0.462994619(GPA_OverALL)

b_1 measures the estimated change in the average value of Y as a result of a one-unit change in X

Here, b_1 = .463 tells us that the average value of a GPA_OverAll increases by .463) = 46.55 ,on average, for each additional GPA_Overall Ratings.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

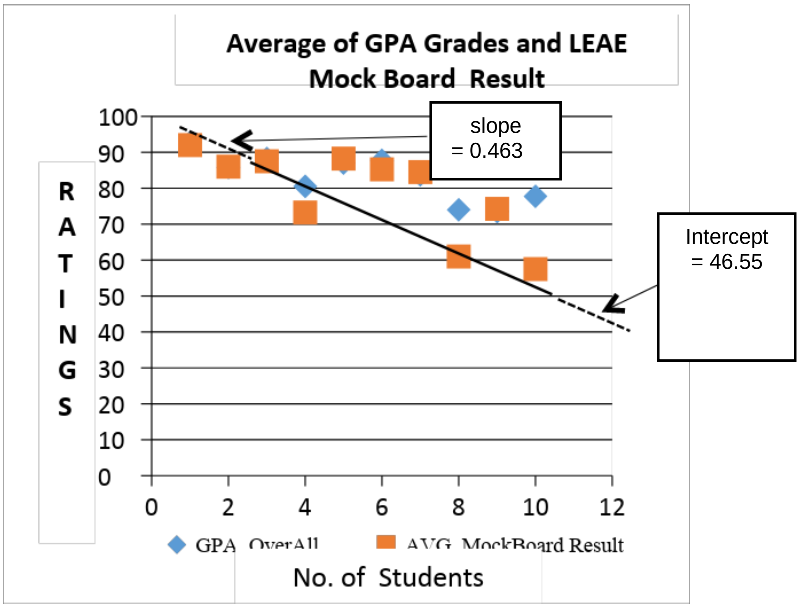


Figure 8. Scatter Plot with intercept and slope

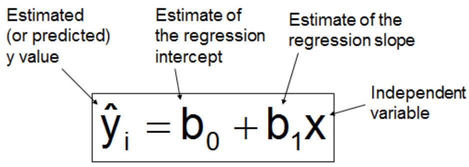
Figure 8 shows the Scatter Plot of x and y variables. As a result Data from Table has slope of 0.463 and an intercept of 46.55.

Sample Prediction:

Predict the Mock Board Exam of a Student with GPA_OverAll of 71 and 91

Does Percentage Ratings of GPA_OveraAll affect its Mock Board Examination Result?

Using Regression Formula, the prediction are as follows:



AVG_MockBoard Result = 46.54657482 + 0.462994619(GPA_OverALL)
With a GPA_OverAll of 71:
AVG_MockBoard Result = 46.54657482 + 0.462994619(71)
AVG_MockBoard Result = 46.55+32.87

Predicted MockBoard Result:
AVG_MockBoard Result = 79.42

With a GPA_OverAll of 91:
AVG_MockBoard Result = 46.54657482 + 0.462994619(91)
AVG_MockBoard Result = 46.55+42.13

Predicted MockBoard Result:

AVG_MockBoard Result = 88.67

As shown in the above result from the Predicted MockBoard Result. The Student who has a GPA_OverAll of 71 will have a predicted Mockboard of 79.42 which means a Moderate Chances in the Licensure Examination will be achieved while student who has the GPA_OverAll 91 will have a predicted Mockboard result of 88.67, which means a Strong Chances to Passed the Licensure Examination will be achieved. Interpretation on the Modeling and Simulation results, as presented by Garciano(2012), the CALRA-DSS will use the Table 8 matrix for Mockboard Average Ratings and the Chances on the Licensure Examination is a model to integrate in predicting the Possible Performance of Students in the Licensure Examination[1].

Table 8. Average on MockBoard Result on chances to LEAE

MockBoard Average Ratings	Chances in the PRC Licensure Examination
91 - 100	Very Strong
80 - 90	Strong
70 - 79	Moderate
50 - 69	Weak
49 below	Very weak

V. CONCLUSION AND RECOMMENDATION

Discussion on this Study concentrated on the Academic Predictors for Licensure Examination in Agricultural Engineering using correlation analysis Pearson-Product Moment Correlation and lienear regression analysis. The Mock Board Questionnaire generated by CALRA-DSS is focus only on the BSAE Program curricular courses provided by the SPAMAST Review Center. Based on the Outcome, the results obtained from the Correlation and Regression Analysis and the attributes obtained from eSMS, the identified Academic Predictors has a strong correlation to Mock Board Examination. A supporting tools developed by Tarun et al. (2014) and Garciano (2012) also revealed that the Mock Board Examination has a strong correlation in predicting the PRC Licensure Examination [6]. If the Academic Predictors and Mock Board Examination is good, then the reviewee is predicted to Passed. Other than the above rules, the reviewee is predicted to fail if students will not focus on the subject’s per area that has a Failed result. It is concluded that with the use of this tool, the ICET Department can implement an intervention program timely before Student would intend to take the LEAE. In general outcome of the study can give a hint to the Students as to which subjects can be considered as predictor variables for their licensure exam scores and hence become their focus of study/review while still studying. Administration can also have an idea based on the correlation results on what to contribute in improving the Agricultural Engineering program. It can help also the Faculty to reevaluate their teaching strategies and subject syllabi that would fit the current need of the Agricultural Engineering students. And lastly, result summary report from CALRA-DSS can serve as a reference in terms of policy-making related to admission and selection as well as retention for the BSAE Program.

REFERENCES

[1] S. Garciano , “ Correlation Analysis Between Academic Variables and Licensure Examination Performance of SPAMAST Graduates.” 2012.
[2] Republic Act No. 8559. An Act Regulating the Practice of Agricultural engineering in the Philippines.
[3] L. Barton, P. Willson, R. Langford, & Schreiner, B. “Standardized predictive testing: Practices, policies, and outcomes”. *Administrative Issues Journal: Connecting* education, practice, and research, 4(2), 68-76.2014.
[4] S. Pachejo, and W. Allaga,. “Academic Predictors of the Licensure Examination for Teachers’ Performance of the Rizal Technological University Teacher Education Graduates”. *International Journal of Educational Research and Technology*. Retrieved from www.soeagra.Original Article com/ijert/ijert.htm.2014
[5] D. T. Quiambao, E. G. Baking, L. M. B. Buenviaje, A. V Nuqui, and R. C. Cruz, “PREDICTORS OF BOARD EXAM PERFORMANCE OF THE,” *vol. 1, no. 1*, pp. 1–4, 2015.
[6] I. M. Tarun, B. D. Gerardo, and B. T. Tanguilig III, “Generating Licensure Examination Performance Models Using PART and JRip Classifiers: A Data Mining Application in Education,” *Int. J. Comput. Commun. Eng.*, vol. 3, no. 3, pp. 202–207, 2014.
[7] BS in Agricultural Engineering in the Philippines, available online: http://www.finduniversity.ph/majors/bs-in-agricultural-engineering-philippines/
[8] H. David, Mannila Heikki, and Smyth Padhraic. “Principles of Data Mining”.
[9] O. Figuerres, “An Analysis of the Performance of the University of Northern Philippines in the Licensure Examination for Teachers’. *International J o u r n an l o f E d u c a t i o n a l R e s e a r c h a n d T e c h n o l o g y* . 2010.
[10] G. Garcia. “Academic Performance as Determinant to Pass the Licensure Examination for Teachers”. *International Journal of Educational Research and Technology*. Retrieved from www.soeagra.Original Article com/ijert/ijert.htm. 2010.
[11] A.,Young, A., & P. Willson, P. “Predicting NCLEX-RN success. CIN: Computers, Informatics, Nursing, 30(1), 55-60.2012.
[12] E. Zweighaft, “HESI Over Time: What the Research Demonstrates and Applications to Improving”. Podium presentation at the 2016 Elevate Outcomes with HESI Conference. Las Vegas, Nevada, 2016.
[13] S. G. Poorman and E. J. Martin, “The role of nonacademic variables in passing the National Council Licensure Examination,” *J. Prof. Nurs.*, vol. 7, no. 1, pp. 25–32, 1991.
[14] M. S. Chen, J. Han, and P. S. Yu, “Data mining: An overview from a database perspective,” *IEEE Trans. Knowl. Data Eng.*, vol. 8, no. 6, pp. 866–883, 1996.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 4, April 2019

[15] V. Ramesh, P. Parkavi, and P. Yasodha, “Performance Analysis of Data Mining Techniques for Placement Chance Prediction,” *Int. J. Sci. Eng. Res.*, vol. 2, no. 8, pp. 1–7, 2011.
[16] S. Yadav, B. Bharadwaj, and S. Pal, “Data mining applications: A comparative study for predicting student’s performance,” *Int. J. Innov. Technol. Creat. Eng.*, vol. 1, no. 12, pp. 13–19, 2012.
[17] C. M. Zahn, A. Saguil, A. R. Artino, T. Dong, G. Ming, J. T. Servey, E. Balog, M. Goldenberg, and S. J. Durning, “Correlation of National Board of Medical Examiners Scores with United States Medical Licensing Examination Step 1 and Step 2 Scores,” *Acad. Med.*, vol. 87, no. 10, pp. 1348–1354, Oct. 2012.
[18] Zupan et al. Orange and Decisions- at-Hand: Bridging Predictive Data Mining and Decision Support. [Online]. Available: http://www-ai.ijs.si/MarkoBohanec/pub/iddmorng.pdf
[19] J. Pascua, and J. Navalta, "Determinants of L.E.T. Performance of the Teacher Education Graduates in a State University". *JPAIR Multidisciplinary Journal*. Volume 6. 2012
[20] B. Duckor, K. Castellano, K. Téllez, D. Wihardini, and M. Wilson, "Examining the Internal Structure Evidence for the Performance Assessment for California Teachers: A Validation Study of the Elementary Literacy Teaching Event for Tier I Teacher Licensure". *Journal ofTeacherEducation*. doi: 10.1177/0022487114542517.2014.
[21] M. Chapple. Regression. [Online]. Available: http://databases.about.com/od/datamining/g/regression.htm
[22] H.L. Lee, C. Park, Agent and Data Mining Based Decision Support System and its Adaption to a New Customer Centric Electronic Commerce, , Expert Systems With Applications, 25, 4, pp. 619-635, 2003