

International Journal of Advances in Electrical Engineering

E-ISSN: 2708-4582

P-ISSN: 2708-4574

IJAEE 2021; 2(1): 36-40

© 2021 IJAEE

www.electricaltechjournal.com

Received: 22-11-2019

Accepted: 27-12-2019

Dr. Adenowo Adetokunbo O
M.Sc. Systems Engineering,
Ph.D. in Engineering,
Department of Electronics &
Computer Engineering, LASU,
Ojo, Lagos, Nigeria

Keshinro Kazeem Kolawole
M.Sc. Computer Science,
Ph.D. in Engineering,
Department of Computer
Engineering, Laspotech,
Ikorodu, Lagos, Nigeria

Green Oluwale A
M.Sc. Computer Engineering,
Ph.D. in Engineering,
Department of Computer Engineering,
Laspotech, Ikorodu, Lagos,
Nigeria

Balogun Wasiu Adebayo
M.Sc., Ph.D. in Engineering,
Research Scholar,
Department of Mechatronics
Engineering, Laspotech,
Ikorodu, Lagos, Nigeria

Correspondence

Keshinro Kazeem Kolawole
M.Sc. Computer Science,
Ph.D. in Engineering,
Department of Computer
Engineering, Laspotech,
Ikorodu, Lagos, Nigeria

Performance evaluation of student academic results using clustering

Dr. Adenowo Adetokunbo O, Keshinro Kazeem Kolawole, Green Oluwale A and Balogun Wasiu Adebayo

Abstract

Student's academic performance is an essential object in the institutions of higher learning for predicting systems of modification that could enhance better academic performance in an institution. Several performance evaluation algorithms have been proposed at Lagos State Polytechnic, South-Western Nigeria, but there is a need to find the best prediction techniques for students' academic progress and performance. This is done by engaging the use of statistical analysis in MINITAB. The students' academic results are examined on a semester-by-semester basis rather than using the overall result or cumulative grade point average (CGPA). This creates clusters in the linear regression graph, which would be used to summarize the performance level at the end of a semester. The clustering of each point to the line of best fit indicates the model is best described by the "linear model" used in this research work for the data analysis. This research work provides a superlative student performance evaluation technique in data mining, as well as the most important student attributes for performance prediction. It further provides student performance enhancement skills with evaluation benefits to students, faculty, and the school management.

Keywords: Clustering, dataset, faculty, performance, results, statistical, students, techniques

1. Introduction

Contrary to the opinion of some authors^[1, 2] that it would be difficult to obtain a detailed performance evaluation of students with average scores except the CGPA may not hold water, as it will hide some details. It will be more accurate to evaluate students', based on their day-to-day performance, which is provided by the average scores. This process will not only be able to consider the students' academic performance but also goes a long way in establishing the extra-curricular gap in the students. Because the best academic performance will be judged based on academic and other activities of a student while on campus.

According to these authors, the CGPA and GPA^[3] are still the best means of evaluating students' academic performance, but they did not realize that even the CGPA, must have been derived from the day-to-day evaluation of some tests, practical's, seminars, quizzes, attendances and by any other evaluation format. The foregoing has not been found to affect the student's ability to maintain a constant CGPA, as they are some of those factors to be considered in implementing improvements in services by management. Hence students' performance evaluation is no doubt a key factor in determining the academic progress^[4] and it will go a long way in determining the contribution needed by the management of the institution to offer services that can result in improvements for the students.

Statistical analysis has been used to evaluate the students' performance, and the clusters generated showed that their performances could be grouped and the conclusion drawn therein.

2. Materials and Method

Statistical analysis was used to model the relations between the independent variable (continuous assessment) and the response variable (total score), based on the "Linear Model" of Statistical Package for the Social Sciences, SPSS 26.0. The ANOVA of the linear regression model of the data analysis showed the level of significance of the model at 96.8%, 94.8%, 87.7%, and 80.3% for the results of CTE 211, CTE 212, CTE 213, CTE 214, respectively (Table 1, 2, 3 and 4). This reveals how well the model fits the results. Therefore, it deduces that the total variance in the "Total score" could be described using this mathematical model. Results from the statistical regression analysis showed the statistical relationship between the independent variable and the response variable.

R^2 is the correlation coefficient, which describes the percentage of “Total results” variation that is described by its relationship with “Continuous Assessment”. Therefore, the R^2 (adjusted) is the percentage of “Total results” variation that is described by its relationship with “Continuous Assessment”, adjusted for the number of predicting variables in the model. This adjustment is significant since R^2 for this model increases when a new independent variable (predictor) is added. Therefore, the adjusted R^2 is a valued instrument for comparing the illustrative power of models with different predictors [5]. The data set has revealed that the examination performance of

the student may or may not depend on the C.A performance, as a student who had poor scores in the C.A., may have good grades in the examination. Conversely, the statistical analysis here showed that there is a linear relationship between the C.A and the total score. Hence, a definite effect will reflect on the final performance of the student for the C.A score. It is therefore expedient for student(s) to prepare better and have good scores in C.A. as it has a direct influence on the final grades.

3. Results

Table 1: Statistical Analysis of CTE 211 ND 2 results

Descriptive Statistics							
	N Statistics	Minimum Statistics	Maximum Statistics	Mean		Std. Deviation Statistics	Variance Statistics
				Statistics	Std. Error		
C.A	50	0.00	36.00	28.2400	1.07649	7.61191	57.941
Total Score	50	0.00	81.00	57.3200	2.46153	17.40585	302.957
Valid N (list wise)	50	0.00					
Model Summary							
Model	R	R Squared	Adjusted R Squared	Std. Error of the Estimate			
1	0.968	0.937	0.936	4.41448			
Anova							
Model		Sum of Squares	df	Mean Squares	F	Sig.	
1	Regression	13909.474	1	13909.474	713.760	0.000 ^b	
	Residual	935.406	48	19.488			
	Total	14844.880	49				
Coefficients							
Model		Coefficients Std. Errors	t	Sig.			
1	Constant	2.422	-2.142	0.037			
	Continuous Assessment	0.083	26.716	0.000			

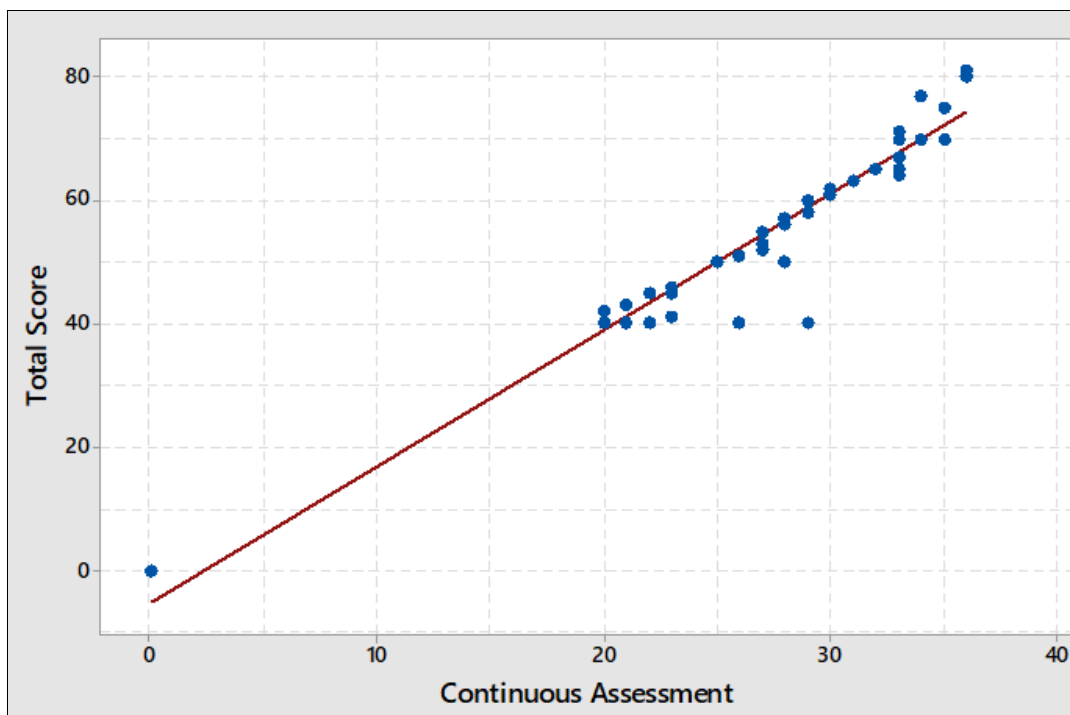


Fig 1: A plot of Total score versus Continuous Assessment for CTE 211 ND 2 results

Table 2: Statistical Analysis of CTE 212 ND 2 results

Descriptive Statistics						
	N Statistics	Minimum Statistics	Maximum Statistics	Mean	Std. Deviation Statistics	Variance Statistics
C.A	50	0.00	78.00	29.9800	6.52308	42.55057
Total Score	50	0.00	78.00	61.8000	14.565660	212.15845
Valid N (list-wise)	50					
Model Summary						
Model	R	R Squared	Adjusted R Squared	Std. Error of the Estimate		
1	0.948 ^a	0.899	0.897	4.66919		
Anova						
Model		Sum of Squares	df	Mean Squares	F	Sig.
1	Regression	9349.217	1	9349.217	428.838	0.000 ^b
	Residual	1046.463	48	21.801		
	Total	1039.680	49			
Coefficients						
Model		Coefficients	Std. Errors	t	Sig.	
1	Constant	3.136		-0.767	4.47	
	Continuous Assessment	0.102		20.708	0.00	

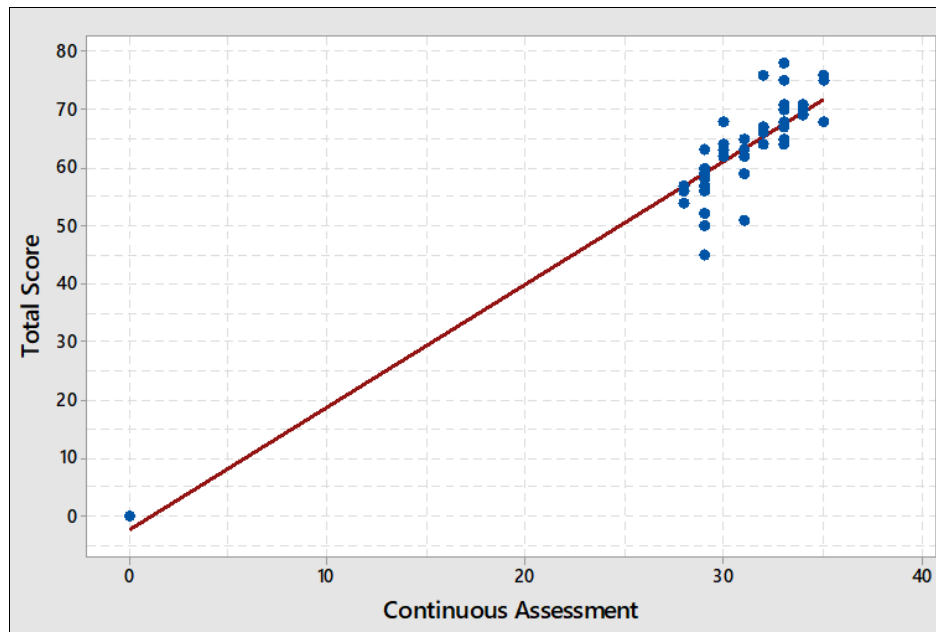


Fig 2: A plot of Total score versus Continuous Assessment for CTE 212 ND 2 results

Table 3: Statistical Analysis of CTE 213 ND 2 results

Descriptive Statistics						
	N Statistics	Minimum Statistics	Maximum Statistics	Mean	Std. Deviation Statistics	Variance Statistics
C.A	50	0.00	38.00	30.7400	6.57394	43.21669
Total Score	50	0.00	66.00	51.7400	13.21689	174.686
Valid N (list-wise)	50					
Model Summary						
Model	R	R Squared	Adjusted R Squared	Std. Error of the Estimate		
1	0.877 ^a	0.770	0.7765	6.40596		
Anova						
Model		Sum of Squares	df	Mean Squares	F	Sig.
1	Regression	6589.878	1	6589.878	160.587	0.00 ^b
	Residual	1969.742	48	41.036		
	Total	8559.620	49			
Coefficients						
Model		Coefficients	Std. Errors	t	Sig.	
1	Constant	0.4374		-0.569	0.572	
	Continuous Assessment	0.139		12.2672	0.000	

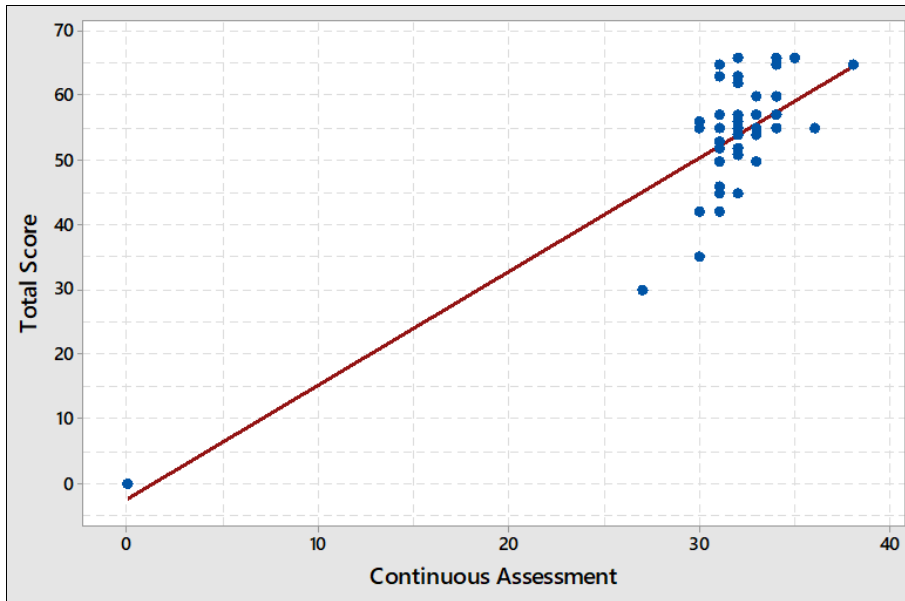


Fig 3: A plot of Total score versus Continuous Assessment for CTE 213 ND 2 results

Table 4: Statistical Analysis of CTE 214 ND 2 results

Descriptive Statistics						
	N Statistics	Minimum Statistics	Maximum Statistics	Mean	Std. Deviation Statistics	Variance Statistics
C.A	50	0.00	28.00	23.7600	5.27435	27.81877
Total Score	50	0.00	82.00	45.0800	15.01814	225.54453
Valid N (list-wise)	50					
Model Summary						
Model	R	R Squared	Adjusted R Squared	Std. Error of the Estimate		
1	0.803 ^a	0.644	0.637	9.04721		
Anova						
Model		Sum of Squares	df	Mean Squares	F	Sig.
1	Regression	7122.782	1	7122.782	87.020	0.000 ^b
	Residual	3928.898	48	81.852		
	Total	11051.680	49			
Coefficients						
Model		Coefficients Std. Errors	t	Sig.		
1	Constant	5.961	-1.549	0.128		
	Continuous Assessment	0.245	9.328	0.000		

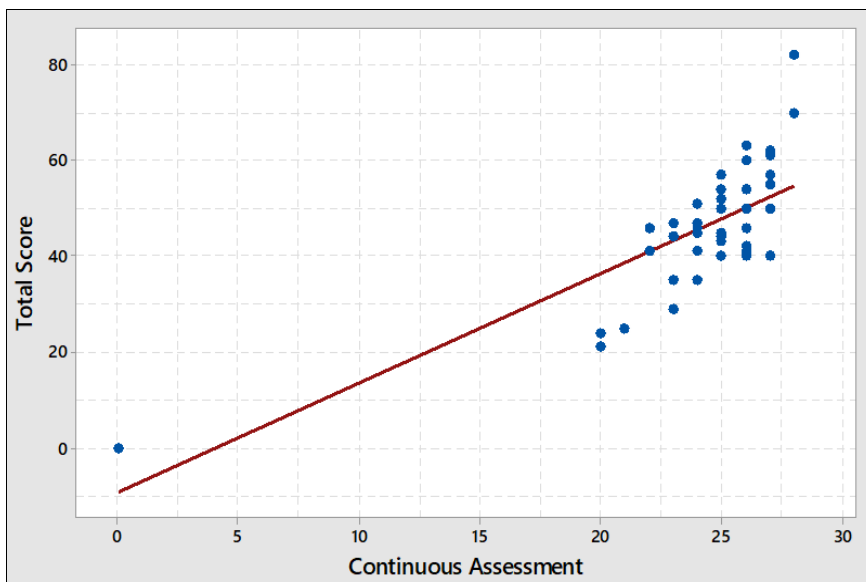


Fig 4: A plot of Total score versus Continuous Assessment for CTE 214 ND 2 results

4. Discussion

Statistical Package for the Social Sciences, SPSS 26.0 was used to assess the descriptive statistics of the data set. The descriptive statistics indicated in tables 1.0 – 4.0 describe the counting statistics for categorical variables and summary statistics for associated numeric variables. The data set presented in this research work is the numeric type, and the total number of observations ($N = 50$) is found to be 50, which indicates reasonable enough data for statistical analysis. The average performance for the whole class in the “continuous assessment” and “total score” is described as the mean values in the tables. Mean values for the continuous assessment (C.A.) are found to be 28.2400, 29.9800, 30.7400, and 23.7600, for CTE 211, CTE 212, CTE 213, and CTE 214, respectively. This shows an increased and better C.A. performance in the order CTE 213 > CTE 212 > CTE 211 > CTE 214. Furthermore, the total score performance is found to be in the performance order, CTE 212 (61.800) > CTE 213 (57.3200) > CTE 211 (51.7400) > CTE 214 (45.0800). This clearly shows that the level of assimilation of an entire class to a particular subject or course work varies with another as seen in the C.A and Total score performance. The variance in each data set or distribution is a measure of dispersion that describes the extent to which the distribution is scattered around its mean. The plot of figures 1.0, 2.0, 3.0, and 4.0 are the linear plot of the total score versus C.A. for CTE 211, CTE 212, CTE 213, and CTE 214, respectively for each student in each course work. The clustering or closeness of each point to the line of best fit indicates the model is best described by the “linear model” used in this research work for the data analysis. Hence, the good grade of students as described as the “total score” strongly depends on the “C.A. performance”.

5. Conclusion

This statistical analysis has been able to show the results cluster and hence gives an insight into the performance of a student in a course of study. This will go a long way in revealing the course performance of each level of the students, thereby enhancing the views of management in terms of service provisioning.

6. References

1. Oyelade OJ, Oladipupo OO, Obagbuwa IC. Application of k-Means Clustering algorithm for prediction of Students' Academic Performance, International Journal of Computer Science and Information Security, (IJCSIS) 2010, 7.
2. Sansgiry SS, Bhosle M, Sail K. Factors that affect academic performance among pharmacy students, American Journal of Pharmaceutical Education 2006.
3. Arbabisarjou Azizollah *et al.* Evaluation Valuation of Academic Performance among Students of Medical Sciences, Int. J Pharm. Technol 8(2), 12241-12247.
4. Yadav RS, Soni AK, Pal S. A Study of Academic Performance Evaluation Using Fuzzy Logic Techniques, Comput. Sustain. Glob. Dev. Int. Conf. on. IEEE 2014, 54-59.
5. Akintola OO, Aderounmu AN, Abiola IO, Abodurin KE, Adeniran T, Agboola F *et al.* Quantitative analysis of phytochemical compounds in barks and leaves of *Okoubaka Aubrevillei* collected from Iwo, South-Western Nigeria', Journal of Bio-residual Management 2020;7(3):131-142.

6. Ramesh Prasad Tharu. Multiple regression model fitted for job satisfaction of employees working in saving and cooperative organization. Int J Stat Appl Math 2019;4(4):43-49.