Journal of the Nigerian Association of Mathematical Physics
Volume33,(January, 2016), pp 99 – 106
© J. of NAMP

Using Discriminant Analysis to Identify Major Prerequisites for Success in Specific Courses of Study in a University System

Iduseri A. and Osemwenkhae J. E.

Department of Mathematics, Faculty of Physical Sciences, University of Benin, P.M.B. 1154, Benin City, Edo State, 300001, Nigeria.

Abstract

Inappropriate choice of course of study at tertiary institutions results in unacceptable levels of attrition and poor throughput rate. The impact on students can cause lasting damage to self-esteem and the consequences can influence an entire lifetime. In this paper discriminant analysis was used to identify major prerequisite for success in a specific course of study in a university system. We present a case study with forty (40) industrial mathematics majors where discriminant analysis was used successfully to determine the major prerequisite for success in industrial mathematics.

Keywords: Major prerequisite, throughput rate, industrial mathematics, discriminant analysis

1.0 Introduction

The economic and social development of a country, as well as the rate of its technological growth is directly linked with students' academic success or achievement. Academic success is an attribute which is not attributed to either man or woman. It is strictly based on the survival of the fittest; the stronger claims its place. High failure rates in specific courses of study [1-4] has not only led to unacceptable level of attrition and poor throughput rate, but the impact can cause lasting damage to self-esteem, and the consequences can influence an entire lifetime. High failure rates are costly to all stakeholders and the University in particular. This is because the throughput of the University is not only reduced, but often results in large class size. According to Heinesen [5], "class size is one of the factors that impact upon academic success and general relationship is a negative one". In addition, majority of the final outputs may result in educated derelicts.

Over the years, research in academic prediction has centered on graduation, withdrawal and selection of the students on the basis of either their collegiate success or cumulative results of Remedial or Pre-ND, and literature to date suggests no loss of interest. A sheer number of these studies used the student GPA to measure the student performance [6-8]. Some other researchers use the result of a particular subject or use the previous year result [9-11].

From the last few years, a number of studies have been carried out to identify and analyse numerous factors that affect academic success in various center of learning [12-14] and a few in specific courses of study in a university system [15, 4]. Majority of these studies focus on different factors such as previous schooling [13], parents' education [16], family income [17], self motivation, age of student, and learning preferences [18], class size [19, 20], extracurricular activities [16], gender difference [21, 20], teacher's education and teaching style [22], and entry qualification [4], Infrastructure provision and learning facilities [12, 14]. The findings of these studies vary from region to region and their results differ in various settings.

While the sheer volume of these studies may be impressive in terms of their utility in providing corrective measures that improve academic performance or guarantee academic success of students especially in public funded institutions, we do perceive a general and skewed interest. Most of the early research and much of the current ones, involved academic prediction in terms of college placement or students' final result with emphasis on the criterion of college success being first session or year grade point average. In a nut shell, Researchers are mostly interested in quality of general input with less emphasis on the quality of general output. Identification of major prerequisite for success in specific courses of study in the area of mathematical sciences has not being given much attention in literature by substantive researchers. In a University system, prerequisite is a course (i.e., a class taught at educational institution as a part of school curriculum that leads to a degree or certificate) for which the understanding of its concept is required as a prior condition for success in any course of study (i.e., a program of study that leads to a degree or certificate from an educational institution).

Corresponding author: Iduseri A., E-mail:augustine.iduseri@uniben.edu, Tel.: +2348036698860

In most developing countries, placement of students into various courses of studies is in part determined by students' preferred course of study. In Nigeria, for a number of institutions, student admission in any course of study is based on whether the student has chosen the course as it first choice. While private funded Universities and a few public funded Universities admit students into their second choice of course of study. However in most cases in public funded Universities, students still found themselves admitted into courses they never choose nor applied for. This is most prevalent in the area of mathematical sciences, in particular are majors such as Pure Mathematics, Applied Mathematics, Industrial Mathematics, etc. Students hardly seek admission into these fields of study or Departments offering them. Therefore, any prospective student who has the requisite entry qualification for these fields of study is admitted irrespective of whether the prospective student chooses the course or not. To further aggravate the problem, the number of public funded universities in Nigeria is grossly inadequate compared to the number of students seeking university education. As a result, most students who gain admission into public funded universities prefers to secure their admission by accepting any course of study giving to them even though is against their choice of study. The implication of such action is that they find themselves in a course of study for which they are unprepared for.

Learning in any course of study is affected by many factors, and a student love for any course of study will boost its learning ability for that course of study. Thus a student, who showed love for particular course of study by chosen it, will be well prepared to face the challenges of the course material compared to a student admitted into a course he or she never liked. In fact, for a student admitted into a course of study he or she never liked, course work is likely to seem grim and difficult, and even meaningless if it is not related to choice.

The purpose of this research was, therefore, to identify major prerequisites for success in Industrial Mathematics as a course of study in a university system. With the rapid growth of banking institutions and financial houses and industries (large and small scales), consultancy (management and engineering) in Nigeria, the application of Industrial Mathematics cannot be over-emphasized. This research aims to equip educators and academic advisors in particular, with requisite knowledge to provide guidance. As a case study, a particular research area was chosen. For confidentiality reasons, the particular research area is called Faculty of Physical Sciences, University X, Edo State, Nigeria. The academic records of some alumni serve as input data for multivariate technique called discriminant analysis (DA).

2.0 Overviewof Background Theory

Discriminant analysis (DA) is a descriptive multivariate technique for analyzing group data. The best known variety of DA is the Fisher's discriminant analysis (or simply linear discriminant analysis, LDA), whose central goal is to describe the differences between the groups in terms of discriminant functions defined as linear combinations of the original variables [23]. The general form of its equation or function is:

$$Z = u_1 X_1 + u_2 X_2 + u_3 X_3 + ... + u_p X_p \tag{1}$$

where Z is the linear discriminant function, u_i are the discriminant weights analogous to $b's$ in the regression equation, and X_i are the predictor variables. DA has two sets of techniques or procedures. These procedures are predictive discriminant analysis (PDA) and descriptive discriminant analysis (DDA). The distinction between PDA and DDA is that, in PDA the focus is on classifying subjects into one of several groups, whereas in DDA the focus is on revealing major differences among the groups [24, 25]. DA conducted for predictive purpose is based on initial set of observations, the group membership of which is known. This discriminant procedure is commonly coupled with an analysis to classify the initial data set. This classification procedure provides information about the quality or accuracy of the predictive model, i.e., the hit-rate or percentage of correct classifications; hence the focus of PDA is prediction as well as accuracy of the hit rates. Precisely, in PDA, the extent of correct classification is of particular interest, whereas in DDA, the function and structure coefficients are the focus, with the hit rate being immaterial [26]. Fundamentally, predictive analysis is possible in many situations where prior designation of groups exists: such as migration/non-migrant status, employed/unemployed, which customers are likely to buy a product or not buy, whether a person is a credit risk or not, e.t.c.

Fisher's discriminant analysis has been used in the fields of education and applied psychology as a "trait-space model" [27, 28], and has proven useful for identifying the structural dimensions along which groups differ with respect to placement problems. Also, LDA has been used to a limited extent in monitoring students' school performance in science [29-31]. These Researchers believed that in tackling the problem of learning diagnosis and instructional placement, educational indicators used should contain cognitive and noncognitive determinants in order to influence performance. To the best of our knowledge, its usefulness in monitoring students' school performance in industrial mathematics with respect to identifying major prerequisite as a cognitive factor capable of influencing performance has not been studied by any substantive researcher.

3.0 Materials and Method

3.1 Defining Grouping Variables for Data Set Collection

For the purpose of groupings, it is a common practice to graduate students with first class (1^{st}), second class upper (2^1), second class lower (2^2) and third class (3^{rd}) based on cumulative grade point average (CGPA). $CGPA_S$, as the average of all of a student's grades for all of his or her complete education career, are considered suitable in designating group memberships. Moreover, the designation procedure is tailored to the long-standing recognition that most students are of average achievement or extremely high, and a very few students have extremely low achievement levels. Consequently, two groups of students in terms of their graduating class of degree were formed: the first group (G1) is students who graduated with first class, second class upper and second class lower honors, while the second group (G2) is students who graduated with third class honors.

3.2 Defining Discriminating Variables for Data Set Collection

In order to obtain relevant predictors of students' school performance with respect to the indicator system mentioned above, we used the student's profile scores for all industrial mathematics courses offered at 200 level in the department of Mathematics, from 2010/2011 to 2012/2013 academic sessions in a University system. These three academic sessions were the sessions for which the Department throughput was significantly reduced when compared to previous session's throughput. In faculties of Mathematical Sciences or Physical Sciences in Nigeria Universities, courses taken by students at 100 level are basically the same across individual courses of study or departments. Courses that constitute an integrated composition within any course of studies are taken from 200 level and above. Hence, all the industrial mathematics 200 level courses were treated as a necessary set of discriminating variables. The number of students for the first group (N_1) is twenty (20), and the number of students for the second group (N_2) is also twenty (20). Therefore, a total of forty (40) students whose group memberships (in terms of graduating class of degree) were established were treated as a training data set with known membership status. Table 1 provides a description of discriminating variables.

Table 1: Description of Discriminating Variables

Variable Name	Description	Credits	Semester
MTH212	Real analysis I	03	1
MTH214	Introduction to operations research	03	1
MTH218	Mathematical methods I	03	1
MTH219	Statistics	03	1
MTH230	Linear algebra	03	1
MTH222	Real analysis II	03	2
MTH227	Introductory numerical analysis	03	2
MTH228	Mathematical methods II	03	2
MTH229	Applied statistical methods	03	2

3.3 Data Analysis

In DA, seeking the subset of best discriminators is considered as a process of mapping the original data into more effective useful discriminating variables [32]. One key way to winnow down a large number of covariates to relatively fewer ones is to employ stepwise methods. A predictive discriminant analysis [33] was performed using the DISCRIMINANT subprogram in SPSS 16 in Windows to determine the important predictors used to distinguish group membership. The METHOD=Stepwise option was chosen to specify the criteria by which the predictor variables would be included in the discriminant analysis. This procedure which results in a forward selection analysis [25], builds the predictive function by entering the predictors, one at a time until the increase in squared canonical correlation (R^2) is no longer statistically significant. This stepwise procedure became necessary because of the number of variables and the fact that there was no unified theory dictating the use of particular variables. However, at the end of the stepwise analysis, we found that MTH218 and MTH219 were the only two variables out the nine potential variables that made significant independent and combined contributions. MTH218 and MTH219 are 200 level courses offered by industrial mathematics students with course descriptions shown in Table 2.

Table 2: Description of MTH218 and MTH219 Courses

COURSE	DESCRIPTION	CREDITS	SEMESTER
MTH218	Some techniques of integration; by substitution by parts and fraction. Differentiation; reduction formula, partial differentiations, applications and classification of critical points of functions of two variables. Lagrangian Multipliers. Coordinate systems; change from Cartesian to polar, spherical and cylindrical coordinate systems. Taylor's and Maclaurin's series. Differential coefficients of the nth order. Liebnitz's rule; application to the solution of differential equations. Complex numbers; Hyperbolic functions, De Moivre's theorem. Roots of complex numbers, Roots of polynomials, Exponential form. Functions of complex variables.	03	1ST
MTH219	Regression and correlation: least square estimation of simple linear regression, interpretation of regression coefficient; use of regression. The product-moment and rank correlation with their interpretation and use. Elementary time-series analysis. Probability: finite sample space, axioms of probability, simple theorems, concepts of probability, addition and multiplication rules, conditional probability and independence, tree diagrams, Bayes' theorem and combinatorial analysis. Probability distributions: random variables, means and variances, Binomial, Hypergeometric, Poisson and normal distributions.	03	1ST

Source: Faculty of Physical Sciences Prospectus (2014/2015 academic session), University X, Edo State, Nigeria.

3.3.1 Building the Predictive Discriminant Function

The problem is to set up a procedure based on the students' grades in MTH218 and MTH219, which enables us to predict the students' correct group when we do not know which of the two groups a student, will likely belong.

Using an arbitrary linear discriminant function (LDF) given by:

$$Z = u_1\left(MTH\,218\right) + u_2\;\left(MTH\,219\right) \tag{2}$$

To determine the vector of raw discriminant weights, u in equation (2), we first compute

(a) Variance-covariance matrices for group 1 and group 2 is given by

$$S_1 = \frac{1}{N_1 - 1}\begin{bmatrix} \sum\left(X_1 - \overline{X}_1\right)^2 & \sum\left(X_1 - \overline{X}_1\right)\left(X_2 - \overline{X}_2\right) \\ \sum\left(X_1 - \overline{X}_1\right)\left(X_2 - \overline{X}_2\right) & \sum\left(X_2 - \overline{X}_2\right)^2 \end{bmatrix} = \begin{bmatrix} 97.696 & 47.064 \\ 47.064 & 222.608 \end{bmatrix}$$

Where X_1 is the vector of scores for the first predictor variable (MTH218) in group 1, $\overline{X}_1$ is the mean of the first predictor variable in group 1, X_2 is the vector of scores for the second predictor variable (MTH219) in group 1, $\overline{X}_2$ is the mean of the second predictor variable in group 1, N_1 is the number of students in group 1

$$S_2 = \frac{1}{N_2 - 1}\begin{bmatrix} \sum\left(X_1 - \overline{X}_1\right)^2 & \sum\left(X_1 - \overline{X}_1\right)\left(X_2 - \overline{X}_2\right) \\ \sum\left(X_1 - \overline{X}_1\right)\left(X_2 - \overline{X}_2\right) & \sum\left(X_2 - \overline{X}_2\right)^2 \end{bmatrix} = \begin{bmatrix} 122.484 & -10.768 \\ -10.768 & 51.042 \end{bmatrix}$$

Where X_1 is the vector of scores for the first predictor variable (MTH218) in group 2, $\overline{X}_1$ is the mean of the first predictor variable in group 2, X_2 is the vector of scores for the second predictor variable (MTH219) in group 2, $\overline{X}_2$ is the mean of the second predictor variable in group 2, N_2 is the number of students in group 2

(b) Pooled sum of squares and cross product matrix (W) is given by

$$W = \left(N_1 - 1\right)S_1 + \left(N_2 - 1\right)S_2 = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} = \begin{bmatrix} 4183.42 & 689.63 \\ 689.63 & 5332.35 \end{bmatrix} \tag{3}$$

where $N_1 = N_2 = 20$ is the number of students for each group

(c) Inverse of matrix, W

$$W^{-1} = \frac{1}{|W|}C$$

$$= \begin{bmatrix} 0.0002442460 & -0.0000315879 \\ -0.0000315879 & 0.0001916198 \end{bmatrix}$$

(d) Mean Vector (d), which is the deviation of group 2 centroid $\left(\overline{X}_{1G1},\ \overline{X}_{2G1}\right)$ from group 1 centroid $\left(\overline{X}_{1G2},\ \overline{X}_{2G2}\right)$ is given by:

$$d = \begin{bmatrix} \overline{X}_{1G1} - \overline{X}_{1G2} \\ \\ \overline{X}_{2G1} - \overline{X}_{2G2} \end{bmatrix} = \begin{bmatrix} 19.6421 \\ 23.1526 \end{bmatrix}$$

So that:

$$u = W^{-1}d = \begin{bmatrix} 0.004066162 \\ 0.003816043 \end{bmatrix} \qquad (4)$$

Thus, substituting these values of raw discriminant weights, u into equation (2), we get a LDF, Z_1 defined as:

$$Z_1 = 0.004066162\ (MTH218) + 0.003816043\ (MTH219) \qquad (5)$$

In order to obtain coefficients that will produce discriminant scores which are measured in standard deviation units, we simply scale the raw discriminant weights. To accomplish this, we first compute the pooled variance-covariance matrix given by

$$S = \left(\frac{1}{N_1 + N_2 - G}\right)W = \begin{bmatrix} 110.425 & 17.366 \\ 17.366 & 137.912 \end{bmatrix}$$

where G is the number of groups and W is the pooled within-groups sums of squares and cross-products matrix (3). The raw coefficient vector (4) becomes

$$v = \frac{u}{\sqrt{u^I Su}} = \begin{bmatrix} 0.061488963 \\ 0.057706635 \end{bmatrix} \qquad (6)$$

where u is the raw coefficient vector (4) and u^T is the transpose of the raw coefficient vector. We therefore redefine the LDF, Z_1 (5) as:

$$Z_2 = 0.061488963\ (MTH218) + 0.057706635\ (MTH219) \qquad (7)$$

This new coefficients, v, is called unstandardized discriminant coefficients or weights (they are called "unstandardized coefficients" because they are to be used with the original data values that have not been standardized). The unstandardized coefficients, v can be used to compare the relative importance of the independent variables in predicting the dependent, much as beta weights are used in multiple regression if the variances of each variable are nearly equal. Should this not occur, the unstandardized coefficients are adjusted by multiplying each by their within group standard deviations. Such adjusted coefficients are called standardized coefficients. Standardized coefficients are the ones that would be obtained if all the predictor variables in the original data had standard deviations of 1.0, which could be achieved by converting the raw data into standard form.

Since the variances of each predictor variable are not nearly equal, the unstandardized coefficients (6) are adjusted by multiplying each by their corresponding predictor variable within group standard deviation. For the training sample, the vector of within group standard deviations for the two predictor variables is given by

$$\sqrt{\frac{w_{ii}}{N-G}} = \begin{bmatrix} 10.49238 \\ 11.84589 \end{bmatrix} \qquad (8)$$

where w_{ii} $(i = 1, 2)$ is the sum of squares for the two predictor variables (3), N is the total number of students for the two groups, and G is the number of groups. Thus the unstandardized coefficient vector (6) becomes

$$v_a = v\sqrt{\frac{w_{ii}}{N-G}} = \begin{bmatrix} 0.64003518 \\ 0.68706162 \end{bmatrix} \qquad (9)$$

Again, we redefine the LDF, Z_2 (7) as:

$$Z_3 = 0.64003518\ (MTH\,218) + 0.68706162\ (MTH\,219) \qquad (10)$$

The estimates of the above standardized canonical discriminant function coefficients indicate the relative importance of the independent variables in predicting the dependent.

4.0 Results and Discussions

Table 3: Canonical variates

Predictor Variables	Raw Coef.	Standardized Coef.	Structure Coef.
MTH218	0.004	0.640	0.775
MTH219	0.004	0.687	0.735

The raw coefficients, standardized coefficients and structure coefficients are shown in Table 3. A cursory look at Table 3 shows that the raw coefficient values, standardized coefficient values and the structure coefficient values are equivalently the same. This implies that the two variables (MTH218 and MTH219) are the same in terms of their order of importance (or unique contribution) in predicting group membership, as well as order of importance by total correlation with the discriminant function.

The confusion matrix, also called a classification table for the analysis of the actual and predicted group membership both for the original data is shown in Tables 4, the rows are the observed categories of the dependent and the columns are the predicted categories, while the percentage of cases on the diagonal is the percentage of correct classification otherwise known as hit rate. The classification results in Table 4 reveal that 97.5% of the students were classified correctly into group 1 or group 2. Since equal sample size will have a 50/50 chance and most researchers would accept a hit rate that is 25% larger than that due to chance, then the overall prediction accuracy of the PDF is significantly better than the expected percent.

Table 4: Confusion matrix for original data

Group			Predicted Group Membership		Total
			1	2	
Original	Count	1	19	1	20
		2	0	20	20
	%	1	95.0	5.0	100.0
		2	0.0	100.0	100.0

97.5% of original grouped cases correctly classified

4.1 Criterion for Performance Evaluation

A criterion for performance evaluation under discriminant analysis is based on misclassification rate, obtained by applying the classification rules derived from the sample to a test data set or to leave-one-out cross-validation samples [34]. Since the training sample size is small, the leave-one-out cross-validation procedure via the training sample was adopted to calculate true error rate.

Table 5: Confusion matrix for cross-validated data

Group			Predicted Group Membership		Total
			1	2	
Original	Count	1	16	4	20
		2	1	19	20
	%	1	80.0	20.0	100.0
		2	5.0	95.0	100.0

87.5% of cross-validated grouped cases correctly classified

The classification results in Table 5 reveal that 87.5% of the students were classified correctly into group 1 or group 2 using the leave-one-out cross-validation procedure. Looking at the values of the overall hit ratio for both the training sample (Table 4) and the cross validated set of data (Table 5), the hit ratio for the cross validated set of data is lower. This is often the case because testing the PDF on the data that gave birth to it is almost certain to overestimate performance. The cross validated set of data is a more honest presentation of the power of the discriminant function than that provided by the original classifications and often produces a poorer outcome. However, since the validated hit rate is more than 25% larger than that due to chance, we therefore conclude that both results are significantly consistent in terms of estimation of the predictive power of the predictive discriminant function.

5.0 Conclusion

The consistency observed in the raw coefficients, standardized coefficients and structure coefficients values for the two predictor variables (Table 3), shows that MTH218 and MTH219 are both effective and useful discriminating variables that have relationship with performance. The significantly high overall hit rate obtained for both the training sample (Table 4) and the cross validated set of data (Table 5) further reveal that MTH218 and MTH219 are the major prerequisite for success in Industrial Mathematics. It also identifies MTH218 and MTH219 as having a booster effect on final graduating Cumulative Grade Point Average (CGPA). In addition, a cursory look at the course description of MTH218 and MTH219 as shown in Table 2 further affirms these two courses as key predictor variables. In fact, the composition of the course outline for these two courses is fundamental with respect to the course objectives.

This means a good understanding of their concepts (i.e., MTH218 and MTH219) is a panacea for success in Industrial Mathematics as a course of study in university X. However, this may not be the case for other universities offering Industrial Mathematics as a course of study. This is because the course description for MTH218 and MTH219 may differ from one university to another. Hence, a different course or courses other than MTH218 and MTH219 may be a panacea for success in Industrial Mathematics for university Z.

The use of discriminant analysis in this manner can be an extremely useful tool for academic advisers, university administrators and policy makers. For university X, a closer analysis of the challenges faced by students that perform poorly in these two courses by academic advisers and university administrators may be worthwhile. A remedial course can be offered to these students (who never choose Industrial Mathematics either as first choice or second choice) as a corrective measure to ensure that they are well-equipped to face the challenges of the course materials. Alternatively, the two courses can further be split to ensure effective teaching and coverage of course outline. Lastly, a standardized multiple measures should be used as a way of assessment.

6.0 References

[1] Anderson, G., and Benjamin, D. (1994). The determinants of success in university introductory economics course. *Journal of Economics Education 25(2), 99-119*.

[2] Zimmer, J., and Fuller, D. (1996). Factors affecting undergraduate performance in statistics: A Review of Literature. Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Tuscaloosa, AL, November 1996). The ERIC Database 1992-2003. ED406424.

[3] Borde, S. F. (1998). Predictors of student academic performance in the introductory marketing course. *Journal of Education for Business 73(5), 302-307*.

[4] Mlambo, V. (2011). An analysis of some factors affecting student academic performance in an introductory biochemistry course at the University of West Indies. *Caribbean Teaching Scholar 1(2), 79-92*

[5] Heinesen, E. (2010). Estimating class-size effects using within-school variation in subject-specific classes. *The Economic Journal 120(545), 737-760.*

[6] Galiher, S. (2006). Understanding the effect of extracurricular involvement. A Research Project Report Presented to the School of Education Indiana University South Bend In Partial Fulfillment of the Requirements for the Degree Master of Education.

[7] Darling, N., Caldwell, L. L., and Smith, R. (2005). Participation in school-based extracurricular activities and adolescent adjustment. [Electronic version]. *Journal of Leisure Research* 37(1), 51-76.

[8] Stephen, L. J. and Schaben, L. A. (2002). The Effect of interscholastic sports participation on academic achievement of middle level school students. *NASSP Bulletin 86, 34-41.*

[9] Hijazi, S. T., and Naqvi, S. M. M. R. (2006). Factors affecting students' performance: A case of private colleges'. *Bangladesh e-journal of Sociology7(2), 34-41.*

[10] Hake, R. (1998). Interactive-engagement vs. traditional engagement "A six-thousand-student survey of mechanics test data for Introductory Physics courses. *American Journal of Physics 66(1), 64-78.*

[11] Tho. (1994). Some evidence on the determinants of student performance in the University of Malaya Introductory Accounting course. *Accounting Education 3(4)*, 331-40.

[12] Mushtag, I and Khan, S. N. (2012). Factors affecting students' academic performance. *Global Journal of Management and Business Research 12(9), 17-22.*

[13] Alam, M. M, Billah, A and Alam, M. S. (2014). Factors affecting academic performance of undergraduate students at International Islamic University Chittagong (IIUC), Bangladesh. *Journal of Education and Practice 5(39), 143-154.*

[14] Sulemana, M., Ngah, I. B. and Majid, M. R. (2014). The perspective of students on factors affecting their academic performance at the tertiary level. *British journal of Education, Society and Behavioural Science 4(8),1021-1028*

[15] Maliki, A. E., Ngban, A. N., and Ibu, J. E. (2009). Analysis of students' performance in junior secondary school mathematics examination in Bayelsa State of Nigeria. *Student Communication Science 3(2), 131-134.*

[16] Islam, M. M. (2014). Factors influencing the academic performance of undergraduate students in Sultan Qaboos University in Oman. *Journal of Emerging Trends in Educational Research and Policy Studies (JETERAPS) 5(4), 396-404*.

[17] Ali, S., Haider, Z., Munir, F., Khan, H., and Ahmed, A. (2013). Factors Contributing to the Students Academic Performance: A Case Study of Islamia University Sub-Campus. *American Journal of Educational Research 1, 283-289.*

[18] Aripin, R., Mahmood, Z., Rohaizad, R., Yeop, U., &Anuar, M. (2008). Students' learning styles and academic performance.22nd Annual SAS Malaysia Forum, 15th July 2008, Kuala Lumpur Convention Center, Kuala Lumpur, Malaysia.

[19] Rodgers, J. R., and Rodgers, J. I. (2003). An investigation into the academic effectiveness of class attendance in an intermediate microeconomic theory class. Faculty of Commerce-Papers, University of Wollongong, N. S. W., 2522 Australia. Accessed on November 12, 2015 from http://ro.uow.edu.au/commpapers/174

[20] Newman-Ford, L., Lloyd, S., and Thomas, S. (2009). An investigation in the effects of gender, prior academic achievement, place of residence, age and attendance on first-year undergraduate attainment. *Journal of Applied Research in Higher Education1(1), 13-28*

[21] Haist, S. A., Wilson, J. F., Elam, C. L., Blue, A. V., and Fosson, S. E. (2000). The effect of gender and age on medical school performance: An important interaction. *Advances in Health Sciences Education 5(3), 197-205.*

[22] Metzler, J. and Woessmann, L. (2010). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. IZA Discussion Paper No. 4999.

[23] Trendafilov, N. T. and Jolliffe, I. T. (2007).DALASS: Variable selection in discriminant analysis via the LASSO. *Computational Statistics and Data Analysis 51, 3718-3736.*

[24] Stevens, J. (1996). Applied multivariate statistics for the social sciences. (3rd ed.)," Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

[25] Huberty, C. J., and Olejnik, O. (2006). Applied manova and discriminant analysis. Hoboken, New Jersey. John Wiley and Sons, Inc.

[26] Onwuegbuzie, A.J., and Daniel, L. G. (2003). Typology of analytical and interpretational errors in quantitative and qualitative educational research. Current Issues in Education [On-line], 6(2). Accessed on November 10 from :http//cie.asu.edu/volume6/number2/index.html.

[27] Cooley, W., and Lohnes, P. (1971). Multivariate data analysis. New York: Wiley

[28] Tatsuoka, M. (1971). Multivariate analysis: Techniques for educational and psychological research. New York: Wiley

[29] Murname, R. J., and Raizen, S. A. (Eds.) (1988). Improving indicators of the quality of science and math education in Grades K-12. Washington, DC: National Press.

[30] Shavelson, R. J., Carey, N. B., and Webb, N. M. (1990). Indicator of science achievement: Options for a powerful policy instrument. Phi Delta Kappan, 692-697

[31] Smith, M. S. (1998). Educational indicators. Phi Delta Kappan, 187-191

[32] Uray, M. (2008).Incremental, Robust, and Efficient Linear Discriminant Analysis Learning. PhD Thesis, Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria.

[33] Huberty, C.J. (1994). Applied discriminant Analysis. New York: Willey and Sons.

[34] Lachenbruch, P. A., and Mickey, M. R. (1968). Estimation of error rates in discriminant analysis. *Technometrics 10, 1-11*