Global Scientific JOURNALS
GSJ: Volume 8, Issue 7, July 2020, Online: ISSN 2320-9186
www.globalscientificjournal.com

SCHOLARSHIP AWARD RECOMMENDER SYSTEM

(CASE STUDY: THE SPRINGTIME DEVELOPMENT FOUNDATION)

Bambi B. J, Ogunlere S. O. Maitanmi O.

Babcock University
Ilisan-Remo
Ogun State
Nigeria

ABSTRACT

The Springtime Development Foundation (SDF), being a non-profit and non-governmental organization has spent over 10 Billion Naira in sponsoring thousands of students who cannot afford the payment of school fees into higher institutions all over Nigeria. However, this quantum of money invested is not commensurable to the performance received as feedback.

From past studies, it is evident that some of the challenges which may have caused this include the method of students' selection. This research then proposes to design a model that will predict and classify students' performance in higher institution putting some factors into consideration.

O'level results, Joint Admission and Matriculation Board (JAMB) exam results and Age of applicant were the factors considered in building this model. A Scikit Learn Neural Network package and a Multi-Layer Perception (MLP) classifier was chosen for the algorithm used for the implementation, while Fuzzy Logic was used as the concept.

From the findings, it was observed that for multifactor/variable system, the traditional binary logic may not be the best logical approach; rather the use of fuzzy logic reduces the complexity and enhances the performance. From the development of the model, it was realized that the complexity of the binary logic will be in order power of (m^n) or (m^{na}), whereas, using the concept of fuzzy logic, the complexity has been reduced to m^a T where 'a' is a constant in both scenario. The output was displayed from the web browser (https://www.sdfscreening.com.ng) thereby hiding all the logic behind the work to the layman user. The output of this model will serve as a determinant for awarding scholarship to the applicants.

From the analysis of the model, it is evident that this model when adopted will be faster than the traditional method since the conventional way of selecting scholarship awardees has not been yielding a positive result in proportion to the students' academic performances in the university, which has resulted to waste of money, time and resources, hence, limit other people chances.

Keywords: Prediction, Academic performance, Fuzzy logic, Algorithm, Neural Network, Scholarship

I INTRODUCTION

The main support of any country is Education and a standard Education could be attained in the study room. (Indriana, Adhistya, Ning, 2013). Poor academic performance of student's in tertiary institutions has become the concern of many, especially in developing countries like Nigeria. One of the factors attributed to Student's poor performance is finance. Young ones (mostly the teenagers could not afford the outrageous bill to go to higher institutions which in turn has led to thuggery, hooliganism, dropped-out, prostitution etc.

Other factors that caused low academic achievement in the higher institutions include but not limited to inadequate textbooks and teaching materials. Teacher factors like lateness to school, frequent absence from work, use of local language in teaching, inability to complete the syllabus, unqualified teachers, low interest in the student's understanding of the subject and lack of dedication to duty.

The Springtime Development Foundation (SDF), being a non-profit and non-governmental organization has deemed it fit to sponsor thousands of students who cannot afford the payment of school fees into higher institutions all over Nigeria, although preferably, Adeleke University Ede. It is therefore generally believed that, once the school fees is paid, students on scholarship have no reason for failure, because; food and accommodation has been taken care of. The university environment is also conducive for learning and competent lecturers are also available. To make students relevant in such a way that, their performance will measure up for the money spent on them is of a great importance to the bodies that fund them such as the Springtime Development Foundation. (Maria, Shade, Nicolae 2015).

The Springtime Development Foundation has spent over 10 Billion Naira in sponsoring students in the university. However, this quantum of money invested is not commensurable to the performance received as feedback. Research shows that some of the challenges that may have caused this include the method of students' selection, which have been traced to be either biasness, subjectivity or inconsistencies. Moreover, if these performances are not improved, may likely make the system collapse.

Therefore, it is highly important to have a foresight of the future performance of students on scholarship. Hence, this research proposes to design a model that predict and classify students' performance in higher institution putting into consideration both the previous academic performance, age and an Intelligence Quotient (IQ) test to find out the authenticity of the previous academic record by adopting predictive algorithms with Fuzzy Logic to serve as a determinant for awarding scholarship to the applicants.

II LITERATURE REVIEW

On the basis of Fuzzy Logic (FL), Ramjeet (2014) proposed a process for students' performance assessment in the university. As an input into the system, the five linguistic variables used were Very poor (VP), Poor (P), Average (A), Good (G) and Very Good (VG), and Excellent (E). In this system, the fuzzy inferences used were 'IF-THEN' fuzzy rules. The downside of this approach is that it requires a complex programming system and can not incorporate various fuzzy platforms.

Maria et al (2015) developed an academic student performance recommendation system that uses a purposeful sampling method with an in-depth technique of interview. The Random tree

algorithm was used in the study because other algorithms performed below random tree. The roptimum algorithm used in the domain of the study is the random tree. The Back-end of the Intelligent Recommender System(IRS) was generated by the rules produced from the optimum algorithm. Hence, informing the students academic performance predictions. Below is a diagram showing the design of the recommendation system.

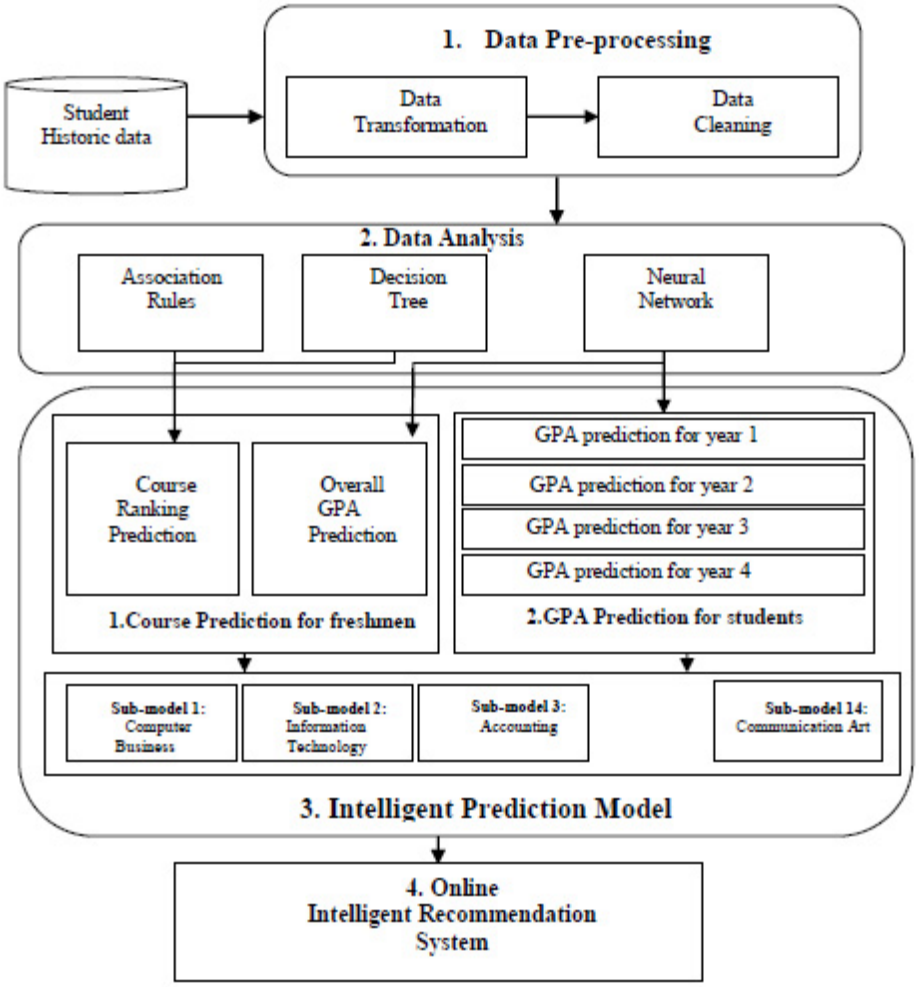


Figure 1, source: Structure of a recommender system, Goga et al (2015)

Izzaamirah(2015) researched on two calculating subjects using the theory of fluctuating logic to identify the quantitative impacts of student performance. The principle applied to 22 First Year students taking two subjects in UNITEN Muadzam Shah. The first two subjects were tested.

Triangular membership feature was the main focus of the fuzzy logic system. Both inputs are sub-X and sub-Y, and the output is performance. The input was analyzed by MATLAB R2014a using a fuzzy logic approach to achieve the finished output. At the end of this study, fuzzy logic was found to be an important means for evaluating student performance. The result shows the output value of two inputs. Of the 22 students 15 were over 0.5 and seven successful students were over 0.75, while 7 were less than 0.5 while none were less than 0.25.

A complete summary of different properties, predictive frameworks, and methods used in educational sector for students' performance prediction was provided by Aljohani (2016), in his research work. It was also evident how important it is to forecast student performance to various parastatals. This is to identify the attrition-risk student early enough to support him and to intervene to increase the retention and rate of performance.

Farshid (2016) compared predictive methods in a course with a standard measurement that defined students at risk of dropping out. The Ensemble model and Naive Bayes Classifiier model with model sequence were found to have the best results among the test modeling methods (such as Supporting Vector Machine, Naive Bayes Classifier and K-Nearest Neighbor).

A model that represents the language of the student, in which the faculty imitates the ability of a teacher to deal with imprecise information was developed by Nidhi (2017), this is shown in Figure 2. The key determinants in identifying successful students are different socio-demographic and study environments. In the estimation and classification of strong moderate and poor students, the Fuzzy ARTMAP network was been used.

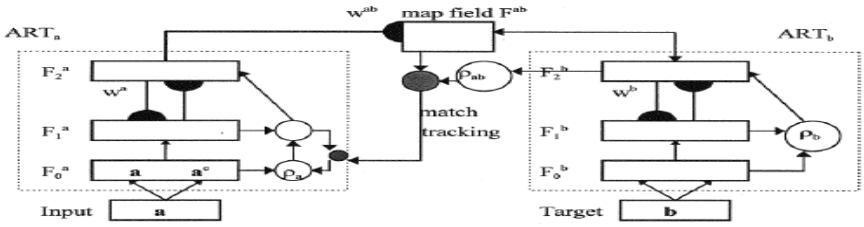


Figure 2, FUZZY ARTMAP ARCHITECHTURE, Source: Nidhi 2017

Suitability of Fuzzy Logic

The weakness recognized above is the essence of applying fuzzy logic. For the system above, the factors can be fuzzified and combined as the aggregate, this average is then categorized and defuzzified to produce the result selection list.

To approach this system with fuzzy logic, weight will be created for each of the factors for example, O level where all the required subject are met such student can be based on aggregate of 5 to 30 (five subject having A1 is 5 and five subject having C6 is 30) this weight can then be inverted as 1 for 6 and 6 for 1.

This weight can also be considered in JAMB looking at for example 180 as baseline for any student and maximum of 400 so these will form the weight sub grouping can also be used like the following:

180-220 as 1

221-260 as 2

Etc

The weight of a student can also be considered for age by inverting age for example

For scale of 1-10

15-17 as 10

18-20 as 2

Etc.

When all these weights are computed and average found then the students can then be grouped into

1. Highly Recommended
2. Recommended
3. Average (Not Certain)
4. Not Recommended
5. Highly not recommended

Using this model, the best of the students will find themselves in the category of highly recommended (i.e. students with aggregate 5 with about 350 – above in JAMB and within 15 - 17).

III Methodology

There are 3 Implementation Modules presented, the first module is the data gathering section where candidates submit their academic application. The second Module is the inference Engine where Fuzzy logic was used as the concept and neural networks for prediction. The third Module interacts back to the user by providing the results of analysis. This research work has been launched on the internet with the link www.sdfscreening.com.ng, showing all the analysis and the system operation.

The frontend was built with Angular framework which consist of HTML, CSS, TYPESCRIPT, The angular framework uses ASP.NET Core framework to interact with the python model used for performance prediction. The angular framework uses Typescript to make API calls to the ASP.NET endpoints and these endpoints make API calls to the prediction model hosted online for responses. Once the response is received, results are stored in a Microsoft SQL Server database for future references.

For the implementation of this work, neural networks algorithm was used. The python scikit learn package was used in the implementation. The python scikit learn package is a package used to perform machine learning operations. It is built off numpy, scipy and pandas. The scikit learn library is open source, light weight and free. The library was mainly selected because of the nature of the project.

Data understanding and preparation

The data was collected from Adeleke University admission unit in raw format. The data contains the following properties:

1. the current level of students,

2. the date of birth,
3. the WAEC scores in each subject,
4. JAMB score in each of the four subjects taken
5. The final JAMB score.

The data was then analyzed, cleaned and processed. Some of the analysis techniques used include correlation, data description, data counting and so on. Missing data was also handled, data normalization was performed (centering and scaling) of the data. Because of the need for uniformity, each student was restricted to having just 8 subjects taken in their WAEC exams. The date of birth was also converted to the age of each student after which the date of birth column was dropped. The scores for WAEC was also converted to numerical values (categorical variables into numerical values) as follows:

A1 = 10

B2 = 8

B3 = 7

C4 = 5

C5 = 4.5

C6 = 4

D7 = 3

E8 = 2

F9 = 0

Below are some of the screenshots before, during and after the data transformation

Out[151]:

	Current Level	Date of Birth	WAEC Results	UTME Scores	UT
55	200	2001-06-12	Economics : C6, English Language : B3, Agricul...	Use of English : 51, physics : 63, chemistry :...	
56	200	2002-10-21	English Language : B3, General Mathematics : A...	Use of English : 58, physics : 47, biology : 5...	
58	200	2001-05-29	Commerce : B3, Accounting : B3, Economics : B3...	Use of English : 56, GOVERNMENT : 64, MATHEMAT...	
59	200	2000-07-31	Economics : C6, Biology : C6, Chemistry : B3, ...	Use of English : 43, PHYSICS : 63, CHEMISTRY :...	
60	200	2002-02-27	General Mathematics : A1, English Language : C...	Use of English : 49, Physics : 51, Biology : 6...	

Fig. 3 Initial WAEC Dataset after separation from main dataset and

Out[135]:

	One	two	three	four
55	Use of English : 51	physics : 63	chemistry : 53	biology : 58
56	Use of English : 58	physics : 47	biology : 53	chemistry : 39
58	Use of English : 56	GOVERNMENT : 64	MATHEMATICS : 68	ECONOMICS : 58
59	Use of English : 43	PHYSICS : 63	CHEMISTRY : 45	BIOLOGY : 55
60	Use of English : 49	Physics : 51	Biology : 62	Chemistry : 54

Fig. 4 JAMB dataset after separation from the main dataset

Out[130]:

	One	two	three	four	five	six	seven	eight	nine
55	Economics : C6	English Language : B3	Agriculture : B3	Chemistry : B2	Civic Education : D7	General Mathematics : B2	Biology : B3	Physics : B3	
56	English Language : B3	General Mathematics : A1	Biology : C4	Chemistry : C5	Physics : C6	Agriculture : A1	Civic Education : C5	Economics : C4	
58	Commerce : B3	Accounting : B3	Economics : B3	Government : A1	Civic Education : A1	English Language : C4	Yoruba : A1	General Mathematics : C5	
59	Economics : C6	Biology : C6	Chemistry : B3	General Mathematics : B2	English Language : B3	Geography : B3	Physics : C6	Civic Education : B2	
60	General Mathematics : A1	English Language : C6	Chemistry : C6	Biology : B3	Physics : B3	Civic Education : B3	Economics : B3	Trade/Entrepreneurship : B2	

Fig 5, WAEC dataset after separation from the main dataset and splitting

Data Modeling

After the data was processed and cleaned, a scikit learn neural network package and a Multi-layer Perception (MLP) classifier was chosen for the algorithm used for the implementation of the model. The mathematical function implemented behind the model is: $f(\cdot) : R^m \rightarrow R^o$ where m is the number of dimensions of input and o is the number of dimensions of output.

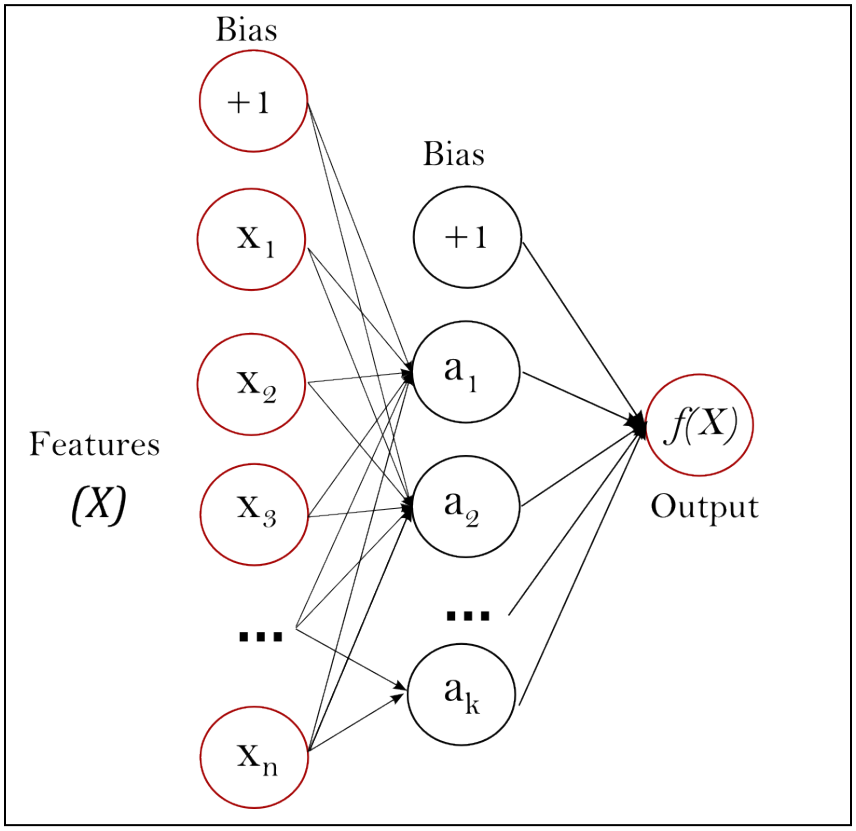


Figure 6, The neural network image representation

Dataset evaluation

Table 1, Model against training dataset

48	0	1	0	0	0
0	1	2	2	0	0
0	0	278	17	0	3
0	0	2	253	0	0
0	0	0	5	0	1
0	0	1	0	0	22

Table 2, Evaluation Matrix for the model using the training dataset

	precision	Recall	F1-score	support
	1.00	0.98	0.99	49
	1.00	0.20	0.33	5
	0.98	0.93	0.96	298
	0.91	0.99	0.95	255
	0.00	0.00	0.00	6
	0.85	0.96	0.90	23
accuracy			0.95	636
macroavg	0.79	0.68	0.69	636
weightavg	0.94	0.95	0.94	636

Table 3, Model against testing dataset

20	1	0	0	0
0	107	11	0	1
0	2	117	0	1
0	0	5	0	0
0	0	0	0	8

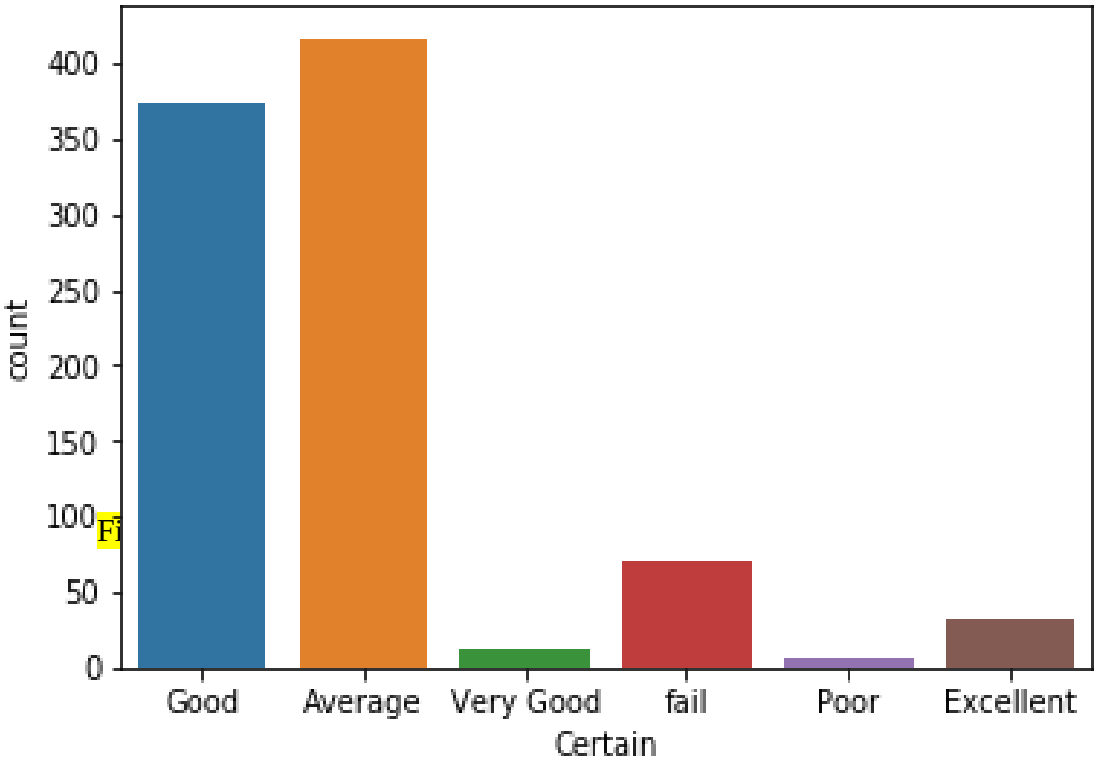
Table 4, Evaluation Matrix for the model using the testing dataset

	precision	recall	F1-score	support
	1.00	0.95	0.98	21
	0.97	0.90	0.93	119
	0.00	0.97	0.92	120
	0.00	0.00	0.00	5
	0.80	1.00	0.89	8
accuracy			0.92	273
macroavg	0.73	0.77	0.74	273
weighedavg	0.91	0.92	0.91	273

Results and Output

The last module of the implementation shows the results of this research work. The results of the model were presented to give the category of students according to their performances. Below is a chart showing the result of the students' according to their academic performances

Performance charts



Discussion of Findings

From the findings, it was observed that for multifactor/variable system, the traditional binary logic may not be the best logical approach; rather the use of fuzzy logic reduces the complexity and enhances the performance. From the development of the model, it was realized that the complexity of the binary logic will be in order power of (m^n) or (m^{na}), whereas, using the

concept of fuzzy logic, the complexity has been reduced to m^a T where 'a' is a constant in both scenario. This shows that with proper knowledge/application of fuzzy logic, will enhance the system performance greatly and also reduces the cost of resources and manpower in the system development. The output was displayed from the web browser (https://www.sdfscreening.com.ng) thereby hiding all the logic behind the work to the layman user. The output of this model will serve as a determinant for awarding scholarship to the applicants.

Recommendations and limitations

With reference to the evaluation done in this study, it is expedient to ensure that Springtime Development Foundation adopt this new modern technique for Scholarship selection since the traditional method of selection has not been yielding a positive result when compared to their academic performances in the university.

This research work has therefore, helped identify any student that might not perform well when they get to the university, it also help the management in decision making by removing biasness from scholarship awardees selection. Although this research met its aim and objectives within the scope of its development, it was limited because the model was designed to function alongside the operation of scholarship selection within Springtime Development Foundation.

Contributions to Knowledge

This study has its unique contribution to knowledge because it proposes an improved method in selecting the right students to be awarded for scholarship in Springtime Development Foundation.

IV Conclusions

From the analysis of the model, it is evident that this model when adopted will be faster than the traditional method since the conventional way of selecting scholarship awardees has not been yielding a positive result in proportion to the students' academic performances in the university, which has resulted to waste of money, time and resources, hence, limit other people chances.

REFERENCES

Aljohani, O. (2016). A comprehensive review of the major studies and theoretical models of student retention in higher education. *Higher Education Studies*, 6(2), 1.

Farshid Marbouti A, Heidi A. Diefes-Dux B, Krishna Madhavan (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Elsivier, Computers & Education* (2016) 1-15

Izzaamirah, B. I. (2015). Application of fuzzy logic to student performance in calculation subjects. *National Symposium & Exhibition on Business & Accounting (NSEBA IV)*, 2-3.

Indriana, H., Adhistya, E. P. & Ning R. (2013). Student Classification for Academic Performance Prediction using Neuro Fuzzy in a Conventional Classroom. *Institute of Electrical and Electronics Engineer* .

Maria G., Shade K., & Nicolae G. (2015). A recommender for improving the student academic performance. *Elsevier* , 1481 – 1488.

Nidhi, A. & Jatinderkumar, S. (2017) Predicting Student Academic Performance using Fuzzy ARTMAP Network, *International Journal of Advances in Engineering Science and Technology.* 187-192.

Ramjeet, S. Y., & Vijendra, P. S. (2014). Modeling Academic Performance Evaluation using Fuzzy C-Means Clustering Techniques. *International Journal of Computer Applications* , 15-23.