Proceedings on Big Data Analytics & Innovation (Peer-Reviewed), Volume 1, 2016, pp.85-96

Students' Academic Performance Modeling and Prediction: A Fuzzy Based Approach

ETUK, S. O., OYEFOLAHAN, I. O., ZUBAIR, H. A., BABAKANO, F. J. & BIMA, M.E.
Federal University of Technology, Minna, Nigeria

ABSTRACT. In higher Institutions of learning, importance is placed on the quality of students admitted as this has direct effect on the quality of graduates been produced by the Institutions, and thus affect the National man-power quality at large. One of the challenges facing the Universities is admitting students on merits and surprisingly, the academic performance of the students admitted on merit begins to drop. Therefore, it is important to predict students' academic performance early enough so as to help instructors take appropriate action in adjusting teaching style and improve greatly on Students' success. In this paper, a fuzzy logic model is used to model data of students and predict their academic performance. Factors like students Ordinary level(O' level) grades, motivation to study in their given course and parents' academic background were used to predict students' academic success level prior to the end of their first academic session. The results when compared with the actual result for the semester examination show 75% accuracy. This early academic performance prediction serves as a guide to the instructor. A good understanding of the students help the instructors to take appropriate steps for effective teaching and learning, and thus improving students' academic performance. Hence, the percentage of students withdrawn from the University after their first academic session due to poor performance (cumulative grade point average of CGPA below 1.5) is expected to be reduced.

Keywords: Student performance, Fuzzy logic, Predictive model

1. Introduction

Education, no doubt is the bedrock on which the development of a Nation lies. A society with quality graduates may bring about higher rates of innovations and faster acquisition of new technology capabilities. Thus, according to (Oladokun, *et al,* 2008), the admission process into any higher institution of learning aims at admitting candidates whose performance would be satisfactory academically in the University. The quality of candidates admitted into any higher institution has a direct impact on the quality of research within the institution and thus has an overall impact on the Nation's development (Oladokun *et al,* 2008). Recently, the decline in the performance of students at the end of their year one in the University is alarming despite that students were all admitted base on merits (Arora and Saini,

2013). Evaluating students' academic performances or achievement using the cumulative grade point average (CGPA) as a pointer is a typical practice in every tertiary academic environment. A percentage of fresh students usually find themselves below the minimum grade point required at the end of their first year (Adedeji, 2001).

This decline can be attributed to several factors affecting effective teaching-learning system amongst which are non- conducive environment of learning and students' lack of motivation to learn (which can be tied to the students' lack of interest in the offered course of study), amongst others. Clearly, the factors considered for admission process is not enough to give teachers a right idea about the level of academic capability of the students they are meant to teach. A very high standard of instruction may not suit a class of more "below-the-average" students while on the other hand, a reasonably lower standard of instruction may be perceived as boring to a class of "more- intelligent-students".

Teachers are considered to be most responsible observers who not only engage class but also monitor the behavior and understanding of students (Arora *et al,* 2013). Hence, if an instructor can properly assess and predict student performance early enough or half way into their first session; then the instructor can take appropriate action to greatly improve the teaching-learning processes geared at improving students' performances. Due to the probabilistic nature of predicting students' performances, fuzzy-based models have been found very effective due to their capabilities to account for fuzzy measures.

In this study, a soft computing approach (fuzzy model)was used for predicting academic performance of students (considering factors like students' O'level grades, UME score, parents' academic background and motivation to learn). The fuzzy model developed was tested using the year one student of the department of Information and Media Technology, Federal University of Technology, Minna as participants.

1. Background of study

A. Student Performance

Academic Background. Students who perform very well in a particular field in an Institution of learning must have had the basic knowledge of the field from their previous learning experiences. In Nigeria, students are grouped into classes of Sciences, Social Sciences and Art while in secondary school based on their performance in related subjects. Consequently, this might translate into the field to which students will be admitted for tertiary education. A student without prior knowledge of the basis of a field or with a weak grade score in the O' level subjects will likely find it challenging to cope with such a field in the University. It is therefore of paramount importance to this research work to use students' academic background as one of the factors in predicting the performance of students in the University.

In previous study, (Adedeji, 2001) used a correlation and regression analysis to investigate the relationship between students' UME scores and their academic performance in the University but due to the nature of UME test as a purely objective

based examination, a combination of Ordinary level academic performance and UME score will provide a better representation of students' prior knowledge.

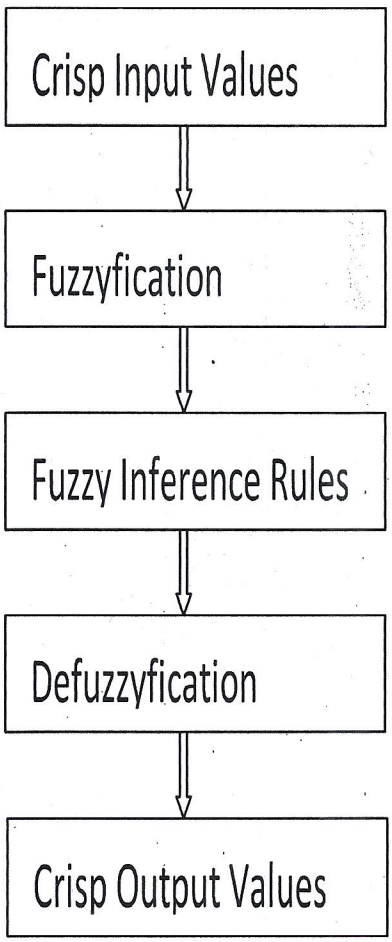
B. Motivation to learn

A driving force to succeed cannot be removed from a success story of any scholar. This also has a very large role to play in students' academic performance. This motivation to succeed academically could be based on factors like interest in a particular course or subject, having a scholar as a role model and wanting to be same, not wanting to disappoint someone who has been part of the students' academic growth, responsibilities of student to alleviate poverty in the family etc. With the right motivation and zeal to succeed, great effort will be put into learning and directly improves academic performance.

C. Fuzzy Logic

Fuzzy logic had long being used to construct better models of reality. Its advantage lies in its ability to provide foundations for approximate reasoning using imprecise propositions base on some fuzzy inference rules. As with sets, fuzzy rules of inference were devised a few decades ago, based on the much older crisp rules. Fuzzy logic shows that truth itself is fuzzy. Rules of inference are rules for deriving truths from stated or proven truths and thus, fuzzy logic are known as efficient tools used to overcome uncertainties related to vagueness, ignorance and imprecision (Yusof *et al,* 2009). Fuzzy Logic involves three main stages namely; fuzzification, rule evaluation and defuzzification. Fuzzification is the process of translating the measured numerical values into fuzzy linguistic values. It is a stage where the degree of membership is determined by applying membership function. Rule evaluation is where knowledge provided by experts is formed, which is then called fuzzy rules (Yusof *et al,* 2009). The fuzzy inference rule will output a fuzzy result, described in terms of degrees of membership of the fuzzy sets.
Defuzzification interprets the membership degrees in the fuzzy sets into a specific action or real-value. This is illustrated in Figure.1

Fuzzy logic has been applied to all fields of life like health to predict cancer, to environment to predict flood detection, and of course, to education to predict students' academic performance.



2. Related Works

A number of works have been done by researchers in an attempt to predict Universities students' academic performance before graduation. This is aimed at having an academic performance assessment of the students to be tutored in order to categorize students and pay more attention on the "very good" students while devising means of supporting the "not-so-good" students.

Existing methods for predicting students' performance include statistical methods. In the case study of (Golding and McNamarah, 2005), they used stepwise multiple regression analysis to predict how factors like students' demographic attribute, qualifications on entry, aptitude test score etc affect the students overall performance. They conclude by suggesting base on their predictions that students with satisfactory predictive performance be allowed to continue their registered program while those with lower performance be channeled to another related program (Golding and McNamarah, 2005).

Other methods researchers used for academic performance prediction is the Data mining algorithms like the decision tree. The work of (Kabakchieva, 2013) used students' personal and pre-University characteristics like gender, birth year,

place of living, total score from previous education etc to predict and classify students into 5 classes which are Excellent, Very good, good, average and bad. They obtained 66.3% accuracy.

Artificial Neural Network (ANN) is another method. (Oladokun *et al,* 2008) developed an Artificial Neural Network (ANN) model considering various factors like ordinary level subjects' scores and subjects' combination, matriculation examination scores, age on admission, parental background, types and location of secondary school attended and gender, among others, to predict the likely performance of a candidate being considered for admission in the University and he achieved an accuracy of 74%. These methods leave a question unanswered: How does it deal with environmental changes and vagueness of reality? This unanswered question lead researchers to using fuzzy based models to predict students' academic performance due to its ability to accommodate uncertainties related to vagueness and imprecision (Osman *et al* 2009) and (Osman *et al,* 2012). In (Yildiz *et al,* 2014), a fuzzy RFM-Model was developed to predict distant learning students' performance. They considered factors like recency, laying importance on the length of time taken before a registered student is admitted on the system; frequency, stating how often an admitted student log on to the system and monetary, considering the length and period of time spent online on the system. The rules of the fuzzy system have been according to expert opinions and the prediction accuracy was 74.7%. The question is; would this model built for distant learning platform generalize to a classroom learning environment?

3. Methodology

The Fuzzy Logic Process

A. Crisp Input Values

In this study, data of year one students of the department of Information and Media Technology, 2015/2016 session were used. Due to the preliminary nature of the study, 36 students in year one were randomly sampled. Out of the 36 questionnaire distributed, 28 were returned, but only 20 responses were valid. This represents 55.60% valid responses. The data for this model was gotten from the students' files and response to the questionnaire on their parents' academic background and what motivates the students to undertake learning in their present course of study. The data were rated "below average", "average" and "good" and a classical fuzzy model was used to predict students' class of degree by predicting their CGPA based on experts opinion.

The crisp input values are the students' data gotten from the students' record and their response to the questionnaire. The process of transforming or normalizing this input values are shown in Table 1.

1. O' level result
2. Motivation to learn
3. Parent's literacy level

S/N	Input Variable		Score	
1	O'level Results	Mathematics	A B C	Good Average "Below average"
		English	A B C	Good Average "Below average"
		Physics	A B C	Good Average "Below average"
		Chemistry / Biology	A B C	Good Average "Below average"
		Geography / Agric science or Economics	A B C	Good Average "Below average"
2	Course Applied for in UME	Present Course Related Not Related	5 – 4 3 1-2	Good Average "Below average"
3	Parent's academic background	Illiterate – pri sch Secondary sch Tertiary Sch	5 – 4 3 1-2	Good Average "Below average"

Table 1: Input Data Transformation

B. Fuzzy Inference Rules

Fuzzy inference rules of the model were created to predict student's academic performance based on expert opinion using O' Level Result, student motivation to undertake their course of study, and their parents Literacy level.

1. If (Result is "below average") and (Motivation is "below average") and (Parent's Literacy is "below average") then (Performance is "below average")
2. If (Result is "below average") and (Motivation is "below average") and (Parent's Literacy is average) then (Performance is "below average")

3. If (Result is "below average") and (Motivation is "below average") and (Parent's Literacy is Good) then (Performance is "below average")
4. If (Result is "below average") and (Motivation is average) and (Parent's Literacy is "below average") then (Performance is average)
5. If (Result is "below average") and (Motivation is average) and (Parent's Literacy is average) then (Performance is average)
6. If (Result is "below average") and (Motivation is average) and (Parent's Literacy is good) then (Performance is average)
7. If (Result is "below average") and (Motivation is good) and (Parent's Literacy is "below average") then (Performance is average)
8. If (Result is "below average") and (Motivation is good) and (Parent's Literacy is average) then (Performance is good)
9. If (Result is "below average") and (Motivation is good) and (Parent's Literacy is good) then (Performance is good)
10. If (Result is average) and (Motivation is "below average") and (Parent's Literacy is "below average") then (Performance is "below average")
11. If (Result is average) and (Motivation is "below average") and (Parent's Literacy is average) then (Performance is average)
12. If (Result is average) and (Motivation is "below average") and (Parent's Literacy is good) then (Performance is average)
13. If (Result is average) and (Motivation is average) and (Parent's Literacy is "below average") then (Performance is average)
14. If (Result is average) and (Motivation is average) and (Parent's Literacy is average) then (Performance is average)
15. If (Result is average) and (Motivation is average) and (Parent's Literacy is good) then (Performance is good)
16. If (Result is average) and (Motivation is good) and (Parent's Literacy is "below average") then (Performance is average)
17. If (Result is average) and (Motivation is good) and (Parent's Literacy is average) then (Performance is good)
18. If (Result is average) and (Motivation is good) and (Parent's Literacy is good) then (Performance is good)
19. If (Result is good) and (Motivation is "below average") and (Parent's Literacy is "below average") then (Performance is "below average")
20. If (Result is good) and (Motivation is "below average") and (Parent's Literacy is average) then (Performance is average)
21. If (Result is good) and (Motivation is "below average") and (Parent's Literacy is good) then (Performance is average)
22. If (Result is good) and (Motivation is average) and (Parent's Literacy is "below average") then (Performance is average)
23. If (Result is good) and (Motivation is average) and (Parent's Literacy is average) then (Performance is average)
24. If (Result is good) and (Motivation is average) and (Parent's Literacy is good) then (Performance is good)
25. If (Result is good) and (Motivation is good) and (Parent's Literacy is "below average") then (Performance is good)

26. If (Result is good) and (Motivation is good) and (Parent's Literacy is average) then (Performance is good)
27. If (Result is good) and (Motivation is good) and (Parent's Literacy is good) then (Performance is good)

C. Crisp Output Values

The Output values illustrated in Table 2 represents the students' Academic Performance based on the University's classification of degrees into '1st Class', '2nd Class Upper', '2nd Class Lower' and '3rd Class'.

S/N	Output Variable	Class	CGPA
1	Good	1st Class – 2nd Class Upper	4.5 – 5.0 3.5 – 4.49
2	Average	2nd Class Lower	2.5 – 3.49
3	"below average"	3rd Class	1.5 -2.49

Table 2: Output Data Transformation

D. The fuzzy based Model

The model below was developed base on students' records to determine their academic performance.

$$\frac{J}{JA_{ve}S} + 0.6\,(Wol1 + Wol2 + \cdots + Wol5) + 0.1(d.f)$$

Where:
J: Jamb score
$JA_{ve}S$: Average Jamb Score.
Wol: Weight of O' Level Result (minmum 5 credit passed courses).
d.f: Dynamic factor.

To account for the dynamic factors, certain questions were asked from the student.

4. Results and Discussion

The Model was tested using feedback from year one students of the department of Information and Media Technology based on the valid responses received. The

responses were used to predict students' CGPA and compared with their actual CGPA as shown in figure 2 below.

The model was 75% accurate in predicting the students' likely class of degree by predicting their CGPA with absolute error displayed in Table 3. This accuracy is of fair performance when compared with existing work in literature of (Oladokun *et al,* 2008) that had 74% accuracy in predicting students' performance.

S/N	Actual CGPA	Predicted CGPA	Absolute Error
1	2.41	2.53	-0.12
2	2.5	3.03	-0.53
3	2.95	2.65	0.3
4	1.5	2.49	-0.99
5	1.05	2.01	-0.96
6	2.77	2.62	0.15
7	2.50	2.69	-0.19
8	1.95	1.7	0.25
9	1.73	2.7	-0.97
10	2.5	2.58	-0.08
11	3.82	3.24	0.58
12	2.64	2.73	-0.09
13	1.55	2.4	-0.85
14	2.91	2.87	0.04
15	3.18	3.17	0.01
16	3.27	3.28	-0.01
17	3.77	3.07	0.7
18	2.05	2.41	-0.36
19	2.77	2.61	0.16
20	2.14	2.43	-0.29

Table 3: Result findings

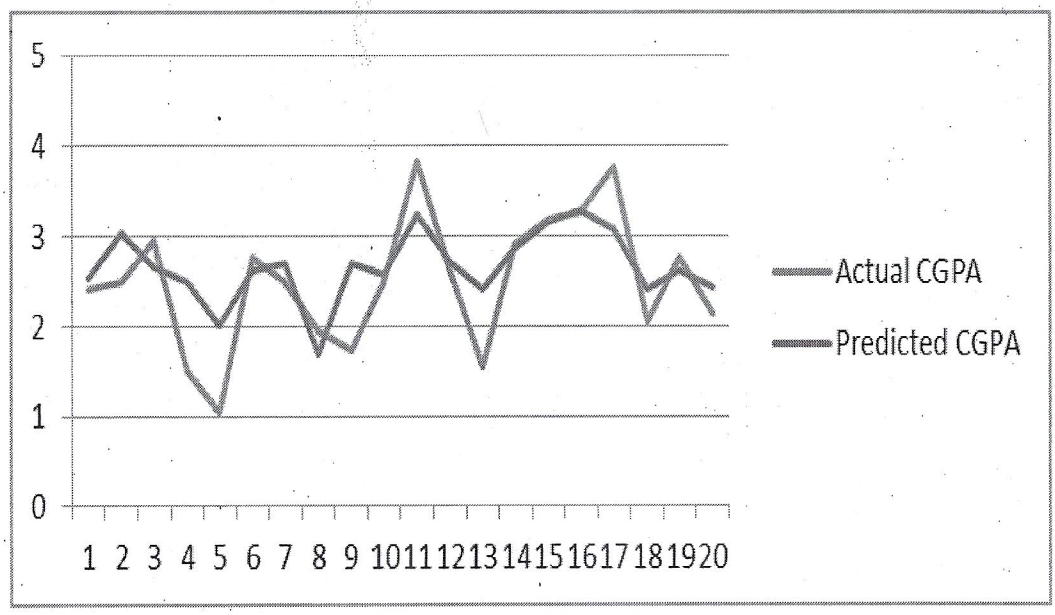


Figure 2: Testing of the Model

Implications of Findings

The result from this study helps the Universities and Instructors in particular to better understand the academic capability of each and every student early enough and adjust the teaching methods so as to better prepare the students to survive the academic pressures. It can be deduced that students whose actual performance is lower than the predicted performance need a special attention in order to live up to the academic requirements. If necessary actions are taken to improve teaching and learning, this will motivates the students to do better. This will have direct positive effect on the percentage of students being withdrawn from the University after their first session for poor academic performance of below 1.5 CGPA. A fuzzy based model was used because of its ability to account for uncertainties and it's solely based on expert opinions. The CGPA predicted are the lower-bound of the degree of class which means the students can achieve the predicted CGPA or higher which is actually fuzzy in nature.

Conclusion

In conclusion, a fuzzy-based mathematical model was developed to predict the academic performance of students'. The test of the model was performed on year one students because it is of paramount importance to begin to understand the students' academic capability from their year one so that all effort can be put in place to assist each students achieve his academic potentials before graduation. The result from the model when compared with the students' 1st semester result shows 75% accuracy and absolute error less than 1.0 in each case. This handy information to the instructors as early as mid session of year one still gives the instructors time to decide on the best teaching techniques for the students. This will improve teach-

ing and learning before the end of the session when students with below 1.5 CGPA are withdrawn.

Limitation of Study

The major limitation of this study is the students' unwillingness to answer and submit the questionnaire.

Future Work

This research study is on-going and an oral interview of students is recommended so as to tackle the limitation of questionnaire and this may likely be an improvement over this model Also, this model can be extended to other department and faculty of the University.

Correspondence
Etuk, S. O.
School of Information and Communication Technology
Federal University of Technology, Minna, Nigeria
Email: abiolastella@futminna.edu.ng.

References

Adedeji, O.B. (2001). "A study of the relationship between students UME results and their undergraduate performance". unpublished

Arora, N., and Saini, J. R. (2013). "A fuzzy probabilistic neural network for students' academic performance prediction". International Journal of Innovative Research in Science, Engineering and Technology, vol 2(9), pp. 4425-4432.

Arora, N., and Saini, J. R. (2013). "Predicting students academic performance using fuzzy ARTMAP network". International Journal of Advances in Engineering Science and Technolody, vol 3(3), pp. 187-192

Golding, P. and McNamarah, S., (2005). "Predicting Academic Performance in the School of Computing & Information Technology (SCIT),‖ Proceedings of 35th ASEE /IEEE Frontiers in Education Conference.

Kabakchieva, D. (2013). "Predicting Student Performance by Using Data Mining Methods for Classification, Cybernetics and Information Technologies, vol. 13, No. 1, pp. 61-72, 2013.

Oladokun, V.O., A.T. Adebanjo, and O.E. Charles-Owaba. (2008). "Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course". Pacific Journal of Science and Technology. 9(1):72-79.

Osman T. and Bahattin K., An adaptive neuro-fuzzy model for prediction of student's academic performance, Computers& Industrial Engineering, No. 57, pp. 732 –741, 2009.

Osman, Y., Abdullah, B., Sevinc, G. and Fulya, D. K. (2012). "A genetic-fuzzy based mathematical model to evaluate the distance education students' academic performance", Procedia-Social and Behavioural Science, vol 55, pp. 409-418

Yildiz, O., Gulsecen, B. And Fulya, D. (2014). "Rules optimization based fuzzy model for predicting distance education students' grades". International Journal of Information and Education Technology, vol 4(1), pp. 369 – 372

Yusof, N., Zin, N. A., Yassin, N. M., and Samsari. (2009). "Evaluation of students' academic performance and learning efficiency based on ANFIS." International Conference of Soft Computing and Pattern Recognition. Pp. 460-465. IEEE Computer Society.