Determination of Academic Performance and Academic Consistency using Fuzzy Logic

Shaikh Diya Salam, Pias Paul, Rehnuma Tabassum, Ifaz Mahmud, Ayat Ullah, Ashiqur Rahman, Rashedur M. Rahman

Department of Electrical Engineering and Computer Science, North South University, Plot-15, Block-B, Dhaka-1229, Bashundhara, Dhaka 1229, Bangladesh.

[shaikh.diya@northsouth.edu](mailto:shaikh.diya@northsouth.edu), [paul.pias@northsouth.edu](mailto:paul.pias@northsouth.edu), [rehnuma.tabassum142@northsouth.edu](mailto:rehnuma.tabassum142@northsouth.edu), [ifaz.mahmud08@northsouth.edu](mailto:ifaz.mahmud08@northsouth.edu), ayatullah7@gmail.com, rashedur.rahman@northsouth.edu

Abstract**— If we want to find a student's performance or consistency, normally we look at his/her CGPA/GPA. But in this new era, this is not the only way to judge a student’s performance and consistency. For our work, at first, we had to find the factors that can be taken into consideration to measure a student's performance and for that we found multiple parameters such as 'Online Assistance', ' Part Time Job', 'Number of Courses', 'Hours of Study' etc. with the help of consultation with several students. Then to obtain data we set up an online survey as well as face to face interaction with the students of Electrical and Computer Engineering students of North South University. Lastly, a system had been built for students, based on the collected data, to find the measurement of performance of a student. Furthermore, it will also help a student find the parameters on which he/she should focus on, in order to improve their performance. In addition to find the performance of students, another system has been built to find the consistency of students with the application of a popular clustering algorithm, i.e., K-Means algorithm.**

*Keywords- Fuzzy logic, Fuzzy inference system, Adaptive Neuro Fuzzy Inference System, K-means algorithm, Academic Performance, Academic Consistency, Factors.*

# Introduction

Research on academic performance is vital for the improvement in academic curriculum of a country, in finding the loop holes and the features which can be added to upgrade the academic system. Additionally, it can help general people understand which factors influence the academic performance of a student.

Academic evolution is one of the best areas for applying fuzzy logic. Since, fuzzy logic provides an output of some certain degree from some imprecise input. Thus, by taking advantage of this property, it can be used to figure out how strong or poor a student’s performance is by taking in imprecise inputs, such as, 'Depression Level', 'Teaching Expectancy' & 'Competency in English' etc. In case of Yadav’s paper [3], it had determined the academic performance of a student was based on two parameters which was not an optimum approach. The reason being, suppose, if a student A scores 50 marks in semester 1 and then score 60 marks in semester 2 and then, there is another student B who has scored 60 in the first semester and 50 in semester 2 then the average score is the same according to the traditional calculation. However, in Yadav’s method he has judged student A’s performance to be improving and student B’s performance to be decreasing since the former student had an increase in score in the second semester and vice versa for the latter. Therefore, with the result of just two semesters such a judgement cannot be authentic.

Parwinder Kaur et al. [1] anticipated a fuzzy system to evaluate students’ performance by considering 5 input variables which are teaching factors, university system, University environment, family factors and personal factors, and further subdivided them into 36 sub-factors. After data collection, the values were mapped into input of fuzzy variable and fuzzy outputs were obtained. They used only FIS to form 243 “if then” rules using knowledge base of the fuzzy expert system and applied Mamdani min implication method to obtain output by matching the user inputs with the fuzzy rules in the knowledge base. Once fuzzy outputs were obtained, they were defuzzified using Centroid method, SOM, LOM and MOM to convert it into crisp set required to obtain single output value.

But as the time changes, the scenario has been changed as well. For example, there was not enough online resources for study purposes back then, now-a-days a student can get enough assistance from online lectures or tutorials which helps in attaining a good grade. In this research, 10 relative parameters have been introduced to determine the academic performance of a student. Moreover, the consistency of students has also been measured, using K-means algorithm by taking 6 different parameters, such as, GPA of S.S.C, GPA of H.S.C, GPA of the last 3 semesters completed and the current CGPA of the students.

The principle focus of this research is to deliver a system to students, where they can evaluate their own academic performance and consistency by varying the range of values for the inputs and find their performance. Thereby, they will be able to comprehend as to which parameters are reflecting upon their performance and thus, determine which parameters they need to focus on for improvement. Moreover, clusters of consistent and inconsistent students have been made from 170 students’ data and, if that student falls into cluster of consistent students and find his performance Low or Average then he can go through our system to look upon which factor or factors are affecting his performance. Even more, if the student desires to perform Good then he/she can evaluate as to which factors that student must concentrate on more. The ANFIS is trained as well with 10 inputs of data. This is done to compare with the proposed method and figure out if the method is authentic.

# Introduction to Fuzzy Techniques

## Introduction to Adaptive Neuro Fuzzy Inference System

ANFIS, also known as adaptive neuro fuzzy system requires the adaptation to learning using fuzzy logic. This is done by means of providing learning environment and inputs to the network which is altered and modified and used by the learning algorithm. The inputs are altered with the help of the learning rule which also influences the learning algorithm. The learning algorithm assesses the size of the sample and the time complexity of the system. The sample size refers to the number of samples to be considered and the time complexity means the speed at which the system will adapt to learning. Even more, the fuzzy logic plays a key role in the adaptation of the learning algorithm by changing the human qualitative knowledge into the process of analyzing the quantitative knowledge. This transformation is not predefined rather it is carried out with the assistance of the learning environment where it adjusts the membership functions in order to learn to adapt to its environment.

In our project, we train our ANFIS with the raw data they gave us and it will generate output membership function based on the inputs using SUGENO implication to test our system.

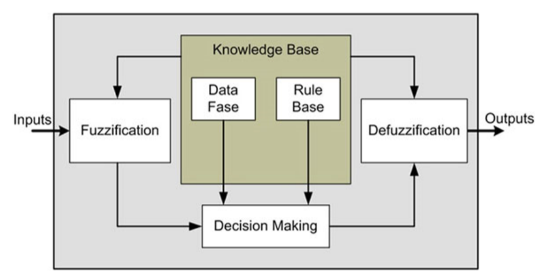
## Introduction to Fuzzy Inference System

The FIS is the abbreviation for Fuzzy Inference System, which is made from three components which are the basic rules. The rules consist of the fuzzy “If-Then” logic rules, the set membership function and the explanation of the fuzzy inference techniques from basic rules to get the output. The FIS works when the actual value of the input is converted to fuzzy values using the fuzzification process through its membership function, where the fuzzy values ranges between 1 and 0.

The knowledge base contains the basic rule and the database, where both these components are the principle elements in decision-making. Usually, the database includes information of the fuzzy set parameter with a function that has been defined for existing linguistic variables. The database is developed by means of defining the universe, determination of the number of linguistic values of the linguistic variable to be used and by establishing the membership functions.

Considering the basic rules, it contains fuzzy logic operators and a conditional statement “If-Then”. The basic rules can be generated from a human or an automatic system, where the searching rules using input-output data numerically. There are several types of FIS like Takagi-Sugeno, Mamdani and Tsukamoto.

However, in our research we have specifically chosen Mamdani technique for rule generation of our proposed method. Below is the detailed structure of the FIS. How we used FIS in our case is briefly discussed in Methodology.



*Figure 1: Structure of the Fuzzy Inference System(FIS)*

## Introduction to K-Means Algorithm

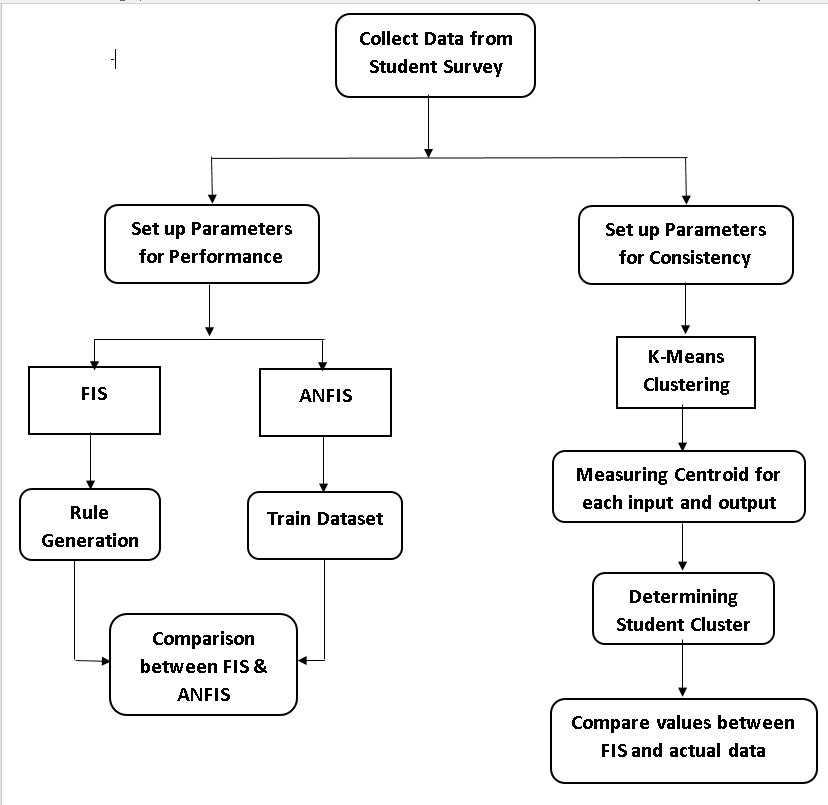
In order, to find a group of data which is similar to one another. Then Clustering algorithm is the best approach. One of the best clustering techniques is the K-means algorithm. It stores *k* centroids, which defines the clusters. A specific point is said to be in a cluster if it is closer to that cluster’s centroid than any other cluster’s centroid. The K-means process is done by alternating between assigning data points to clusters based on the present centroid and choosing centroids (points which are the center of a cluster) based on the current assignment of data points to clusters.

In the research paper 'TV series recommendations from Movie data' [6], they used K-means to make clusters of users movie consumption ratio and given rating for a specific genre to find out the preferability of that genre.

For example, if a user watches movies mostly from the comedy genre, and had given highest ratings to comedy movies then their system will suggest that user series that fall into comedy genre the most. In this case, we used the data from the six parameters mention in Fig.4 to make the clusters of consistent and inconsistent students. And if a student is very consistent then what his performance in our method. If his performance fall into HIGH or VERY HIGH range then we can say our system is giving accurate prediction.

# Research Methodology

To determine the academic performance and to check consistency of the students, different steps were taken. Figure 2 will give a basic step by step overview about our work.



*Figure 2: Flow chart of the Research Methodology*

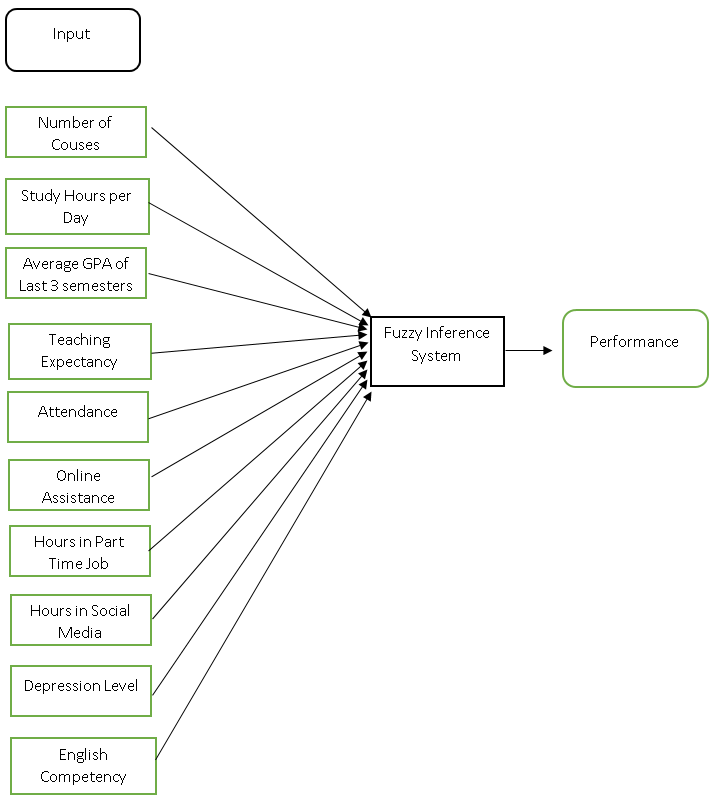
## Data Collection

Data collection was required for finding the factors or parameters which could influence the academic performance on the students. Since, the research is done on university students for which the data was collected on the students from ECE department of North South University (NSU). The data had been collected in two common ways for 170 students. First, the data was collected by face to face interaction with students who had completed at least 3 semesters. Secondly, the other method was by means of online question survey where the answers to each of the question were given in a range of multiple choices from which students could have good flexibility in choosing an option.

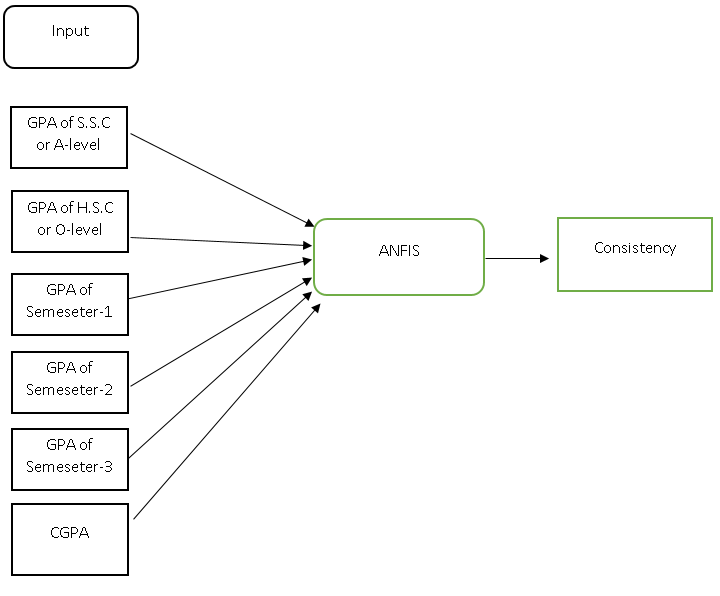
In case of academic performance, multiple factors, shown in Figure 3, had been taken into account, namely, the average attendance of students, the online assistance that students had taken regarding their academic courses, the average number of courses taken by students per semester, the average hours spent in social media and in part time jobs, the number of hours put into studying, the teaching expectation of students, the competency in English of the students, the depression level and the current CGPA of the students. The different factors have been taken into account and they were each given different degrees by means of some value which were named as Very High, High, Medium, Low and Very Low mentioned in Table-2.

For academic consistency, Figure 4, the SSC/O’Level, HSC/A’Level, the GPA of each of the last three semesters and the current CGPA of the student have been considered.

After the completion of the survey, the outputs and results of each student was recorded in an excel sheet. In the excel sheet, each column indicated the parameters or attributes and the rows held the values of each attributes



*Figure 3: Factors taken as input for Evaluation of Academic Performance*



*Figure 4: Mapping inputs for Consistency*

## Data Analysis

This step is a follow up of the data collection process. From this process, we identify the factors that have an impact on the student performance and the academic consistency of the student and we try to determine how each factor varies.

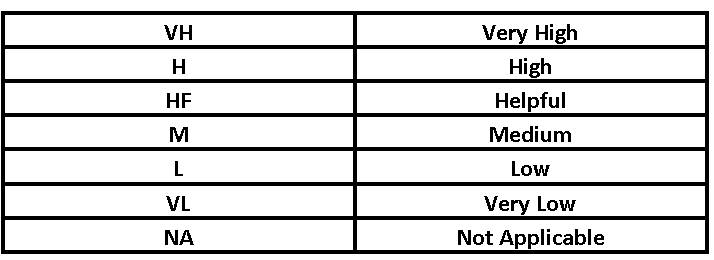
The factors or parameters which have been taken into consideration for evaluating the academic performance had gone through the process of data analysis. The number of courses taken in a semester by a student does have an effect on their grade as it has a direct relation with the pressure of studies and the number of hours for studying increases proportionately. Such parameters also have an impact on the average GPA of the past 3 semesters completed. Moving on, attendance per class, English competency of a student as to how well the student can deliver their work in English and understand the faculty’s lecture also can reflect on their class performance. When students fail to understand a certain topic in class then they can take assistance from the online tutorials which help them in understanding a certain topic. Hours in part job include the time spent in taking private tuition classes of intermediate and elementary level students, which is quite a common scenario in Dhaka for the undergraduate students, running online businesses or freelancing, to earn pocket money. Hours in social media include time spent behind social media like Facebook, Snapchat, Instagram or any other social media applications. The factor, teaching expectancy, means whether the teacher’s method of teaching met the expectation of the student of the courses which they had done in the last three semesters. Lastly, the depression level considers the stress level of a student which occurs in a student due to immense pressure of studies, family issues, romantic or emotional relationships, etc. which affects the mental state of a person, thereby, the student loses focus and concentration on their academics. Such factors have a direct impact on the result as they all consume time and energy which are the essential items for studying. By pointing in these factors for determining the academic performance, can give a thorough and accurate result.

Since, it is a norm to evaluate the performance of a student by means of their academic result, in other word CGPA of an undergraduate student. That is why, the academic performance has been scaled from ‘0.0’ to ‘4.0’ according to the CGPA scale.

In case of assessing the academic consistency of a student, fewer factors have been examined in comparison to evaluating the academic performance. This is because the mentioned factors are good enough to assess the consistency. The consistency of academic performance from the ordinary or secondary school level, which is from grade 10 to their present status in student life. The parameters include, the GPA of SSC/Ordinary Level, GPA of Higher Secondary school or Advanced level, each of the GPA of the last three semesters completed and, the current CGPA of the student. After running K-Means algorithm the clusters of different level of consistent students is found. We tried to take some other factors in consideration to determine consistency such as the number of credits passed so far, how many GED courses had been taken for the last 3 semesters. But, it is found that if Student-A completed 100 credits and has CGPA 3 and Student-B has completed 50 credit and has CGPA 3, level of consistency is not the same for both the student. Since, student B’s CGPA may fall down after 100 credits. Thereby, that may increase the amount of error in our calculation. To maintain the accuracy of our dataset we took data from 3rd and 4th year students.

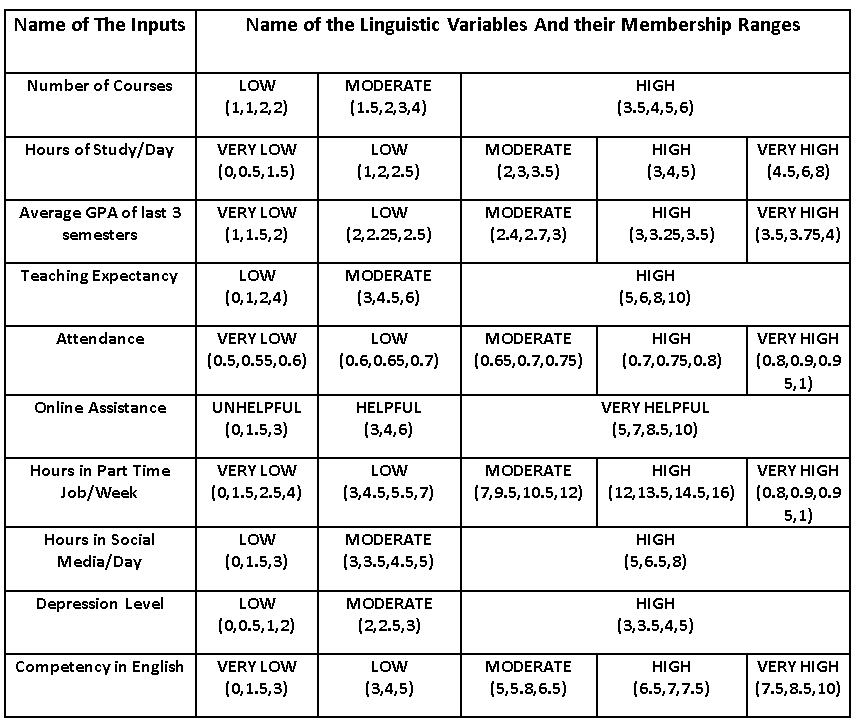
## Rule Formation & FIS for Academic Performance

Table 1: Linguistic Variables



In this step, to build the fuzzy knowledge base mentioned in Figure-1 we generate 50 if-then rules using our membership values for each inputs (Table-2) considering all the real life scenarios depending on our inputs parameters that a student may face that can affect his performance.

Table 2: Linguistic Variables and their membership values for FIS



The rule formation had taken multiple linguistic variables as given in Table 2 upon which we thinks at what case a student can perform in Figure-5.

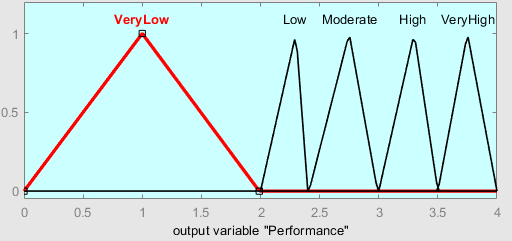


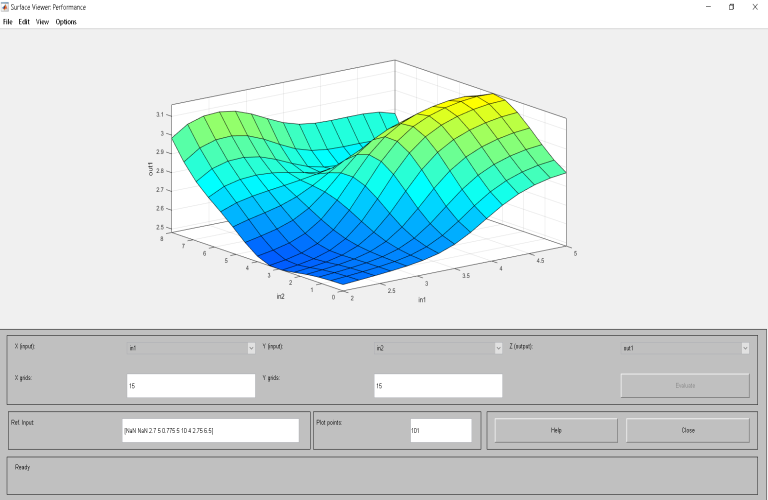
Figure-5: Output membership function in FIS

For the process of rule formation, real life cases were taken into account. Suppose,

1. If number of courses were L, the number of study hours is L, average GPA is L, Teaching Expectation is AV, attendance is M, online assistance is HF, Hours spent behind part-time job is M, hours spent behind social media is L, the depression level is L, the English competency is M THEN performance is L
2. If number of courses were L, the number of study hours is M, average GPA is L, Teaching Expectation is M, attendance is H, online assistance is HF, Hours spent behind part-time job is VL, hours spent behind social media is L, the depression level is L, the English competency is M THEN performance is M
3. If number of courses were L, the number of study hours is H, average GPA is M, Teaching Expectation is M, attendance is H, online assistance is VH, Hours spent behind part-time job is VL, hours spent behind social media is L, the depression level is L, the English competency is M THEN performance is H
4. If number of courses were L, the number of study hours is VH, average GPA is H, Teaching Expectation is H, attendance is VH, online assistance is HF, Hours spent behind part-time job is L, hours spent behind social media is M, the depression level is M, the English competency is H THEN performance is H
5. If number of courses were L, the number of study hours is VL, average GPA is VH, Teaching Expectation is H, attendance is H, online assistance is VH, Hours spent behind part-time job is M, hours spent behind social media is M, the depression level is L, the English competency is VH THEN performance is VH
6. If number of courses were L, the number of study hours is VH, average GPA is H, Teaching Expectation is H, attendance is VH, online assistance is VH, Hours spent behind part-time job is L, hours spent behind social media is L, the depression level is M, the English competency is H THEN performance is VH
7. If number of courses were L, the number of study hours is M, average GPA is M, Teaching Expectation is L, attendance is M, online assistance is NA, Hours spent behind part-time job is M, hours spent behind social media is M, the depression level is M, the English competency is M THEN performance is L
8. If number of courses were L, the number of study hours is M, average GPA is M, Teaching Expectation is M, attendance is VH, online assistance is HF, Hours spent behind part-time job is VL, hours spent behind social media is M, the depression level is H, the English competency is M THEN performance is M
9. If number of courses were L, the number of study hours is H, average GPA is H, Teaching Expectation is M, attendance is H, online assistance is VH, Hours spent behind part-time job is VL, hours spent behind social media is M, the depression level is M, the English competency is M THEN performance is VH
10. If number of courses were L, the number of study hours is VH, average GPA is H, Teaching Expectation is H, attendance is H, online assistance is HF, Hours spent behind part-time job is M, hours spent behind social media is M, the depression level is L, the English competency is M THEN performance is VH

The result from the application of FIS is the determination of the student performance based in the rules loaded into the system in the FIS editor. With the assistance of the rules, it produces the academic performance. In order to find the authenticity of the performance, it is then compared with actual CGPA of the student.

As for the FIS application, 50 rules had been generated which were added into the FIS editor. The result of the rules had been seen in a surface view manner (*Figure-6*).



*Figure 6: Surface View for FIS on Academic Performance*

## ANFIS for Academic Performance

In the ANFIS editor, 170 dataset were loaded where it contained 10 inputs. Based on this dataset, the ANFIS tool generated 170 rules and an output.



*Figure 7: Rule View of ANFIS for Academic Performance*

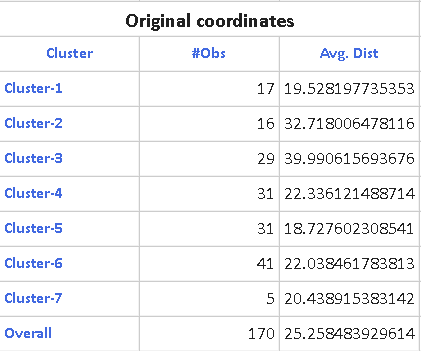
Figure-7 represents output of Performance of 170 students. With this we will compare our proposed method’s output to justify our prediction that we have made for our Fuzzy Rule Base.

## K-Means Algorithm applied for Consistency

The correct choice of k in K-Means is often ambiguous. In addition, increasing the value of k without penalty will always reduce the amount of error in the resulting clustering. To the extreme case of zero error each data point is considered its own cluster (i.e. when k equals the number of data points n). In Equation-1[7] they used 'Mean Squared Error' formula to determine the value of k for making cluster of students with good result, fair result, very good result based on their one semester grades.

In our case, we used 'Elbow Method' to determine the optimum value of k. This method looks at the percentage of variance explained as a function of the number of clusters. More precisely, if one plots the percentage of variance explained by the clusters against the number of clusters, the first clusters will add much information, but at some point the marginal gain will drop, which will give an angle if we draw it in a graph. This "elbow" cannot always be unambiguously identified. We find that in our case after k=7 variance doesn't vary at greater extent. So we made 7 clusters of 170 student’s academic result showed in Table-3.

Table 3: Data for the K-Means Algorithm



Here in column no 2 we can see after 50 iterations how many students belongs to cluster-1,cluster-2 and so on.In row-3 the average distance represents distances between in clusters centroid.

# Result analysis

As our goal was to make a system where a student can evaluate himself, membership functions have been made for fuzzification and then the rules have been generated to measure the performance.

## Validation of the FIS output

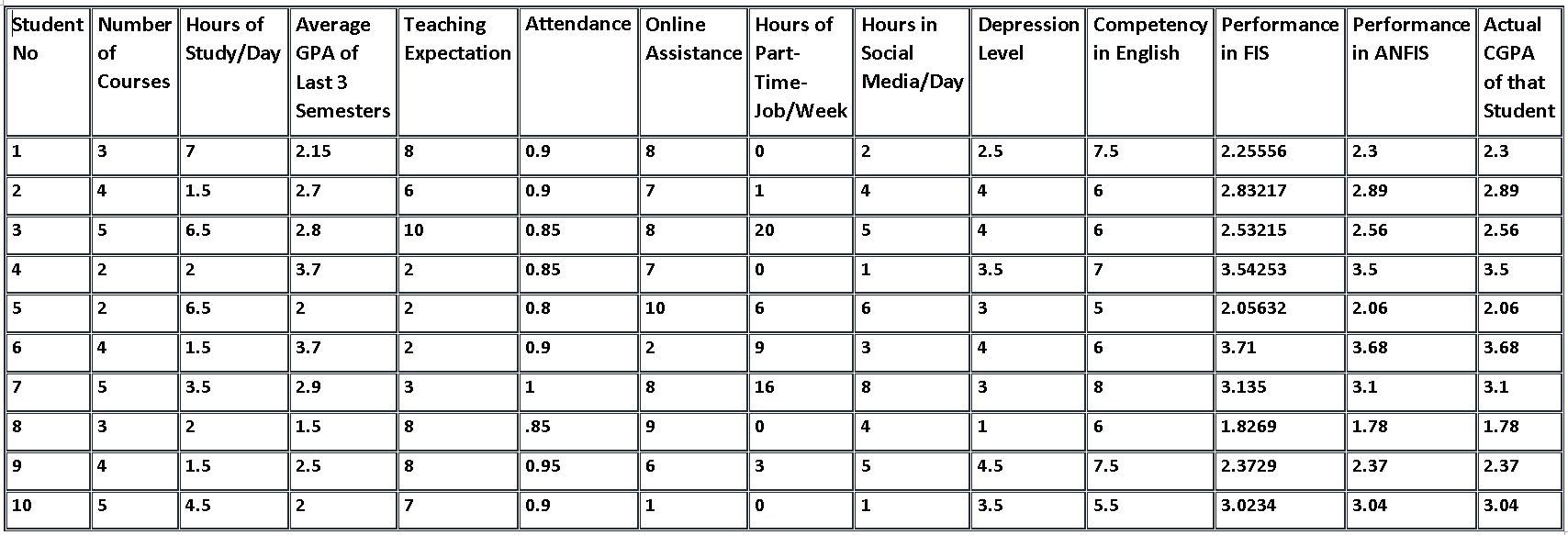
So, if a student found his performance to be below Average or below Good he can go through the proposed method to identify the factors that he should focus on to improve his performance.

For instance, if a student A has CGPA 2.4 which is assigned as low in our method. Then he can go through our system to analyze his performance and based on his average GPA of last 3 semesters and current CGPA he can decide that should he take LOW numbers of courses, assess how many hours he should study per day, whether he should work on any part time job, how many he should spend on social media and, if he doesn't get enough help from teachers then should he look for online resources or spend more time on other sources of study material. In addition, they can examine the factors 'Depression Level’ and 'Competency in English', to observe to what degree these factors are affecting their performance.

## Comparison between ANFIS and FIS on academic Performance

For the comparison of the ANFIS and FIS application on determining the academic performance of a student, the data that had been taken and analyzed were loaded into the ANFIS editor. In the ANFIS editor, the training file had contained the data on which sub-clustering had taken place. After 8 epoch (iterations) on the data loaded it had given an error of 0.000001. From this error value, it can be concluded that the training of the ANFIS is optimum and the output that will result from the ANFIS will be almost accurate.

Table 4: Comparison of FIS & ANFIS



As we see in Table-4, there are data of 10 inputs from 10 students from which we note, both their output in FIS & ANFIS and compare it with their actual CGPA.

It is found that, as ANFIS generates in the output SUGENO implication, the output is computationally efficient. On the other hand, FIS is generating the output using MAMDANI-Min implication. That’s why the output is intuitive. But we got our FIS output result as expected.

Such as, Student-8 in Table-4 is already in probation means he has a ‘Average GPA of last 3 semesters’ of 1.5, but he took 3 courses in that semester, although he got enough help from teachers and online resources but, due to lack of his study hours per day and because he had taken Moderate number of courses, his performance did not improve that much which is seen from his ‘Actual CGPA’. Our FIS system also yields similar output from its knowledge base. Therefore, if that student has a desire to improve his performance then he can decide by changing other input’s ranges to see which combination of inputs values matches his expected performance.

We could have used Sugeno for our FIS rule generation. However, as we know Mamdani is more suitable for human-input values that’s why we used FIS for our system.

Now, we will find the clusters of different types of consistent & inconsistent students and compare that if a student is very consistent then what his performance in our FIS method and actually what his performance from his CGPA.

## Finding the Consistency

After applying K -means clustering with value of K=7, from Table-5 we found 17 & 31 number of ‘Very consistent’ students in cluster 1&4 accordingly. In cluster 2 & 6 we found 16 & 41 numbers of consistent students. In cluster no 3, 5&7 we found 29, 31 & 5 numbers of inconsistent students.

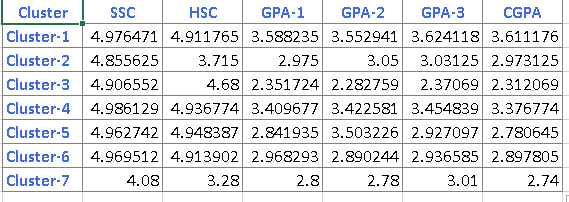


Table 5: Clusters of different types of consistent students

Table-5 represents centroid for each input provided. Now, if there is a student’s GPA of semester-1 then we can measure the Euclidean distance of that variable from each of the other cluster’s centroid of GPA-1 and then, place that student’s GPA-1 in the cluster where we get minimum distance. From this process, it can be determined as to which cluster a certain student will belongs to.

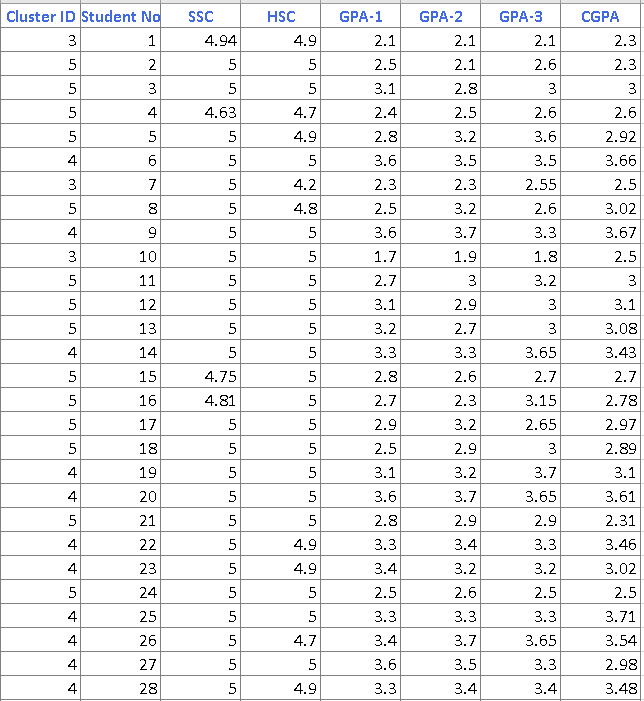
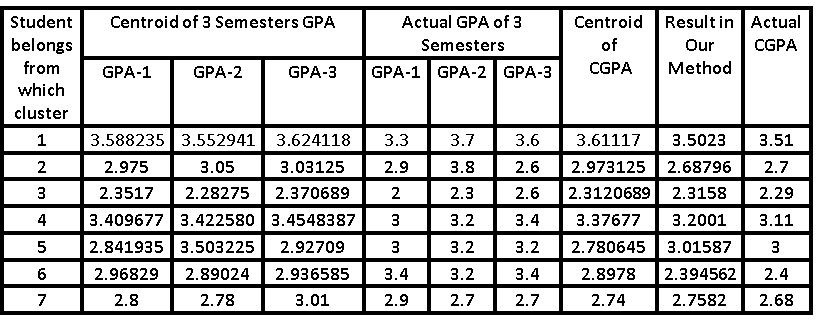


Table 6: Showing Clusters of each individual Student

Table-6 represents the cluster to which each student belongs to which has been obtained from running the K-Means algorithm on the academic results of 170 students. To give an example, we have previously determined cluster 3, 5 and 7 belongs to inconsistent students. In Table -6 student no-15 has a GPA of 4.75,5, 2.8,2.6,2.7 in S.S.C , H.S.C , SEM-1, SEM-2, SEM-3 respectively and CGPA 2.7. Thereby, that student should be found in any of the inconsistent clusters. In Table-6 we can find that Student No-15 belongs to cluster no-5 which we have already determined after implementing K-Means that it belongs to inconsistent students.

Now, we will pick random student data from different clusters to check if a student was found in very consistent or consistent or inconsistent cluster then what will be his performance in FIS to find out our method is working well to give prediction.

Table-7: Validating our FIS by comparing consistent and inconsistent students’ data



From Table-6 we found student number-145 belongs to cluster no-1, which we labelled as “Very Consistent”. As a result, his performance should be ‘HIGH’ or ‘VERY HIGH’ in our method. Eventually, the inputs of that student had been loaded in our method in our system and it was found that the output came to be 3.5023 which falls into “VERY HIGH” range in our method showed in Figure-5. The result has also been compared with the actual CGPA of the student to check the authenticity.   
  
 Similarly, student number-75 from Table-6 has these following results in his S.S.C ,H.S.C, GPA in 3rd last semester, GPA in 2nd last semester and last semester respectively 4.94,4.47,1.9,1.73,1.9. And his current CGPA is 1.98. In our cluster analysis he falls into cluster no-3 which we defined before as “Inconsistent” cluster by analyzing centroid values for all the inputs in Table No-5. Now we entered his given values for our FIS method. Though his number of courses were LOW ,Attendance were MODERATE , didn’t have any part time job but his other parameters didn’t fall on parts that’s why his CGPA was 1.93 and his performance in our method was 1.9523 which we delimited for Very\_Low performance .

Thus, also from this approach, it is found that our proposed system is useful as its outcome equals the expected result.

# Conclusion and Future Work

We made a two-way approach using 12 factors to make sure our proposed Fuzzy Inference System is predicting a student’s performance accurately. Though we tried to cover all the useful parameters, some other parameters such as Relationship Status, Parental approach could also be considered. But as the students did not want to disclose such personal matters.

Although, this system evaluates a student based on our specified education system. However, it can be used on any other university or college student.

##### References

1. P. kaur, P. Agarwal, S.K. Singh and L. Jain “Fuzzy Rule Based Students’ Performance Analysis”, IEEE, Interational Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT) ,2014
2. K. A. Rasmani and Q. Shen, “Data-driven fuzzy rule generation and its application for student  
   academic performance evaluation”, Springer(2006)
3. R. S. Yadav, A.K. Soni and S.Pal, “A Study of Academic Performance Evaluation Using Fuzzy Logic Technique”, International Conference on Computing for Sustainable Global Development (INDIACom),2014
4. O. A. Molina and S. Ratte**,”** Students and teacher academic evaluation perceptions:Methodology to construct a representation basedon actionable knowledge discovery framework”, Spinger,18th February 2016
5. S. Pavani, P. V. S. S. Gangadhar and K.K.Gulhare, “Evaluation of Teacher’s Performance using Fuzzy Logic Techniques”, International Journal of Computer Trends and Technology, 2012
6. Muyeed Ahmed, Anirudha Paul, Mir Tahsin Imtiaz, Md. Zahid Hassan, Shawon Ashraf, Rashedur M Rahman,"TV Series Recommendation Using Fuzzy Inference System, K-Means Clustering and Adaptive Neuro Fuzzy Inference System"IEEE, 2017
7. Oyelade,O.J,Oladipupo,O,Obagbuwa, I. C,"Application of k-Means Clustering algorithm for prediction of Students’ Academic Performance",International Journal of Computer Science and Information Security,2010
8. Abdur Rashid Khan, Hafeez Ullah Amin, Zia Ur Rehman,

“Application of Expert System with Fuzzy Logic in

Teacher’s Performance Evaluation”, International Journal of Advanced Computer Science and Applications, February 2011