Batch: SC_1	**Roll No.: 1911027**
Experiment No. 10	
Grade: AA / AB / BB / BC / CC / CD /DD	
Signature of the Staff In-charge with date	

Title: Case study on Fuzzy systems

Objective: To estimate flat rates using fuzzy logic.

Expected Outcome of Experiment:

CO2 : Apply basics of fuzzy logic.

Books/ Journals/ Websites referred:

1) **https://link.springer.com/chapter/10.1007/978-3-319-59614-3_2**
2) **https://link.springer.com/chapter/10.1007/978-3-319-57261-1_32**
3) **https://fuzzysystems.org/**
4) **Fuzzy Logic with Engineering Applications by Timothy J Ross**
5) **J.S.R.Jang, C.T.Sun and E.Mizutani, "Neuro-Fuzzy and Soft Computing", PHI, 2004, Pearson Education 2004.**
6) **Davis E.Goldberg, "Genetic Algorithms: Search, Optimization and Machine Learning", Addison Wesley, N.Y., 1989.**
7) **S. Rajasekaran and G.A.V.Pai, "Neural Networks, Fuzzy Logic and Genetic Algorithms", PHI, 2003.**

Pre Lab/ Prior Concepts:

1. **Basics of fuzzy systems**
2. **Flat rent estimation techniques**

Implementation Details:

Abstract:
The main objective of this research is to develop a system based on fuzzy logic to estimate flat rent. A number of factors influence the decision of a tenant to rent a flat. Since it is not desirable to work with a large number of variables, we used one of the dimensionality reduction techniques, known as Principle Component Analysis (PCA). PCA helps in reducing the number of variables from the set of data, as well keeping hold the most of the variability in data. This paper describes the implementation of an adaptive neuro-fuzzy inference system (ANFIS) -based approach to estimate the rent of the flat. The Sugeno ANFIS model is proposed in order to develop a systematized approach of generating fuzzy rules and membership function parameters for fuzzy sets from a given set of input and output data.

Introduction:
Fuzzy logic is a computing approach that allows the value of the variable in the interval [0,1] whereas in the Crisp logic the value can be either 0 or 1. In 1965, Lotfi A Zadah introduced fuzzy sets where a more flexible use of membership is possible. In fuzzy sets, many degrees of memberships are allowed. To control and model uncertain system in industry or for any real life applications fuzzy logic is a powerful mathematical tool. The main goal of this paper is to develop a system based on fuzzy logic where the flat rent of any apartment can be estimated. In this paper, the variables used for data analysis are collected from flats in different blocks Residential Area. The data is collected through surveys, from which we could also determine different preferences. The rent of a flat is dependent on many variables. In their survey they have included 14 variables; 5 out of them are crisp, 5 are fuzzy and the remaining 4 variables can be answered by yes/no. Then Principal component analysis (PCA) is applied to reduce the number of inputs and then fed to the ANFIS model. PCA is a dimensionality reduction technique that reduces the number of inputs. It is especially useful for simplifying the interpretation of highly multivariate data sets by reducing the number of variables. It still captures most of the variability of the original data set. The output from ANFIS predicts the estimated rent. The Sugeno ANFIS model presented in this paper was proposed to develop a systematic approach for generating fuzzy rules and membership function parameters for fuzzy sets from a given input – output data set.

Data Set Description

We used 14 variables in our research. Those are discussed below:
1) Distance from Main Road: The far the apartment from the main road, the lower the rent of the flat is.
2) Size of the flat: It is needless to say that the flat size has a direct influence on the flat rent.
3) Rent: It is needed for generating estimation rules which is the main objective of the research.
4) Floor: People give preference in the location of floor. The higher the floor is located, the lower the price is. Generally it is noticed that the 2nd floor is the most preferable.

5) Number of Bedrooms and Bathrooms: It is also important variable that influence the rent. It also dependents on personal preferences.
6) Transportation Availability: The higher the transportation availability, the more preferable the flat is.
7) Lift, Garage, Generator: Availability of these facilities definitely causes increase in the flat rent.
8) Security Level: The higher the security level is, the more preference the flat gets regardless the type of area or tenant.
9) Road condition, Water & Gas supply, Ventilation & Light: These are some environmental parameters that have more or less effect on a particular flat rent.
10) Furnished by Flat Owner: This variable has an effect since full furnished flats can be preferable to students who rent a flat until their end of academic year.

The data set consists of 63 individual flats in different blocks of Bashundhara residential area, Bangladesh. From the total of 63 data sets, 27 were collected through
door-to-door survey and remaining 36 were collected using online survey questionnaire. Initially 52 online surveys were filled up, but due to several errors such as inconsistent information, blank answer space and so on, 16 data sets were excluded.
One sample data set is given in table 1.

Table 1. Sample data set

Variables	Sample Data after transformation
Distance from main road (km)	1.8
Size (Square ft.)	2100
Rent (Tk) (Output)	30,000
Floor	5
Number of Bedrooms	3
Number of Bathrooms	4
Transportation Availability	8
Lift	1
Furnished by the Landlord	1
Road condition	9
Security	8
Garage	1
Generator	1
Water & Gas Supply	5
Ventilation & Light	8

It could be seen from table 1 that distance from main road, size, floor, number of bedrooms and bathrooms are five variables that can take crisp inputs. From the remaining variables, transportation availability, road condition, security, water & gas supply and ventilation and light are fuzzy variables. Finally, the availability of the facilities of lift, furnished by the landlord, garage and generator can be answered only by saying 'YES' or 'NO'.

Methodology

Rent estimation system described in paper has two main parts which are PCA and ANFIS.

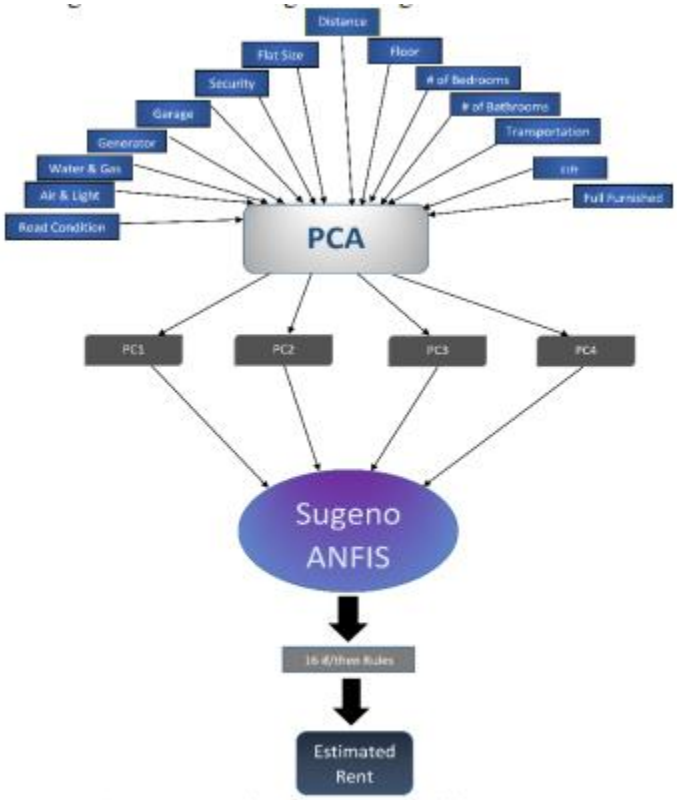
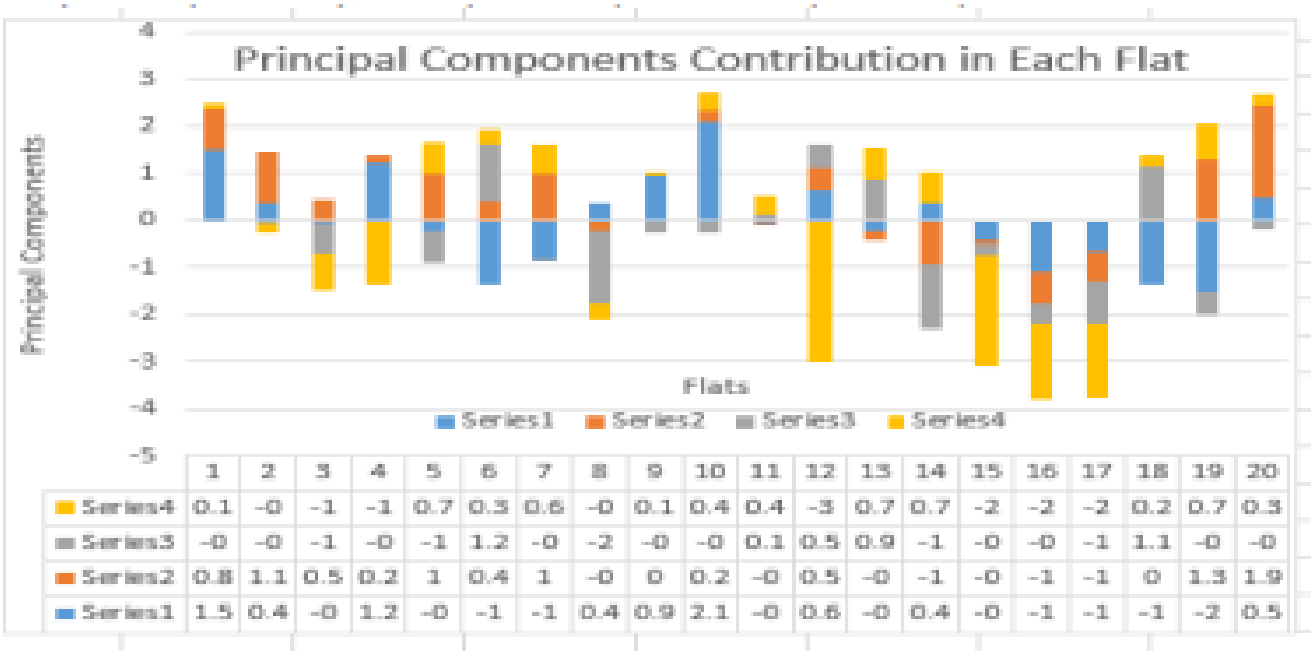


Fig 1. Rent estimation system architecture

PCA: For Principal Component Analysis, the values of all variables need to be in crisp form. The 5 variables with crisp value were taken as they are, the 5 fuzzy variables values are taken as a rating from 1 to 10, where 1 is worst and 10 is best. For the remaining 4 variables, the taken value is 0 if the answer is 'NO' and 1 if the answer is 'YES'. The sample data input is given in table 1 for each case. The 14 inputs are re- duced to 4 principal components using princomp(data matrix) function in MATLAB R2016b. The PCA analysis on our data matrix generates the coefficients matrix, i.e., the matrix of data values transformed into the principal component space and the vector containing the eigenvalues [3]. From the generated eigenvalues, it is observed that the 1 st four principal components capture 99.9% of the total variability in the data set [5]. Figure 2 shows the contribution of the 4 principal components in the data set

where series1, series2, series3 and series4 represent PC1, PC2, PC3 and PC4 respectively. In horizontal axis 20 flats of the 63 are represented. There are 20 stacked columns along the vertical axis that are the Principal component contributions of PC1(series 1), PC2(series 2), PC3(series 3) and PC4(series 4).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Series4	0.1	-0	-1	-1	0.7	0.3	0.6	-0	0.1	0.4	0.4	-3	0.7	0.7	-2	-2	-2	0.2	0.7	0.3
Series3	-0	-0	-1	-0	-1	1.2	-0	-2	-0	-0	0.1	0.5	0.9	-1	-0	-0	-1	1.1	-0	-0
Series2	0.8	1.1	0.5	0.2	1	0.4	1	-0	0	0.2	-0	0.5	-0	-1	-0	-1	-1	0	1.3	1.9
Series1	1.5	0.4	-0	1.2	-0	-1	-1	0.4	0.9	2.1	-0	0.6	-0	0.4	-0	-1	-1	-1	-2	0.5

Fig 2. Stacked column s of Principal Components in each Flat

The outputs of the PCA are the 4 Principal Components for each of the 63 cases. One sample output is shown in table 2.

Table 2. Sample output of PCA

Principal Components	Values
PC1	1.225080245
PC2	0.155955533
PC3	-0.027694021
PC4	-1.346951764

ANFIS: The Sugeno ANFIS model is selected for estimating an approximation of flat rent. Using PCA, now the input of this ANFIS has become the 4 Principal Components PC1, PC2, PC3 and PC4 and the corresponding flat rent will be the output. After training the data and generating FIS, the structure of the ANFIS is as shown in Figure 3

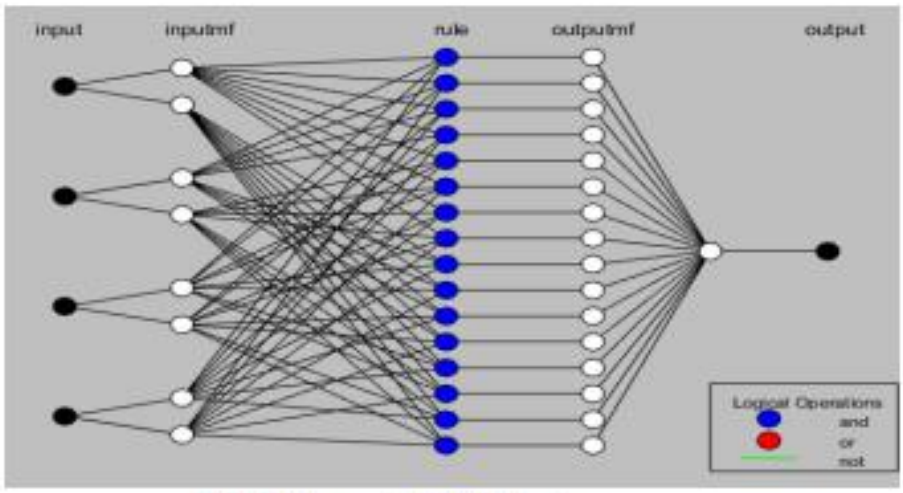


Fig 3. Generated ANFIS structure

The generated rules by Sugeno type ANFIS are given in TABLE 4. Here each variable values are fuzzy numbers with different ranges that correspond to different grade of memberships.
result and analysis

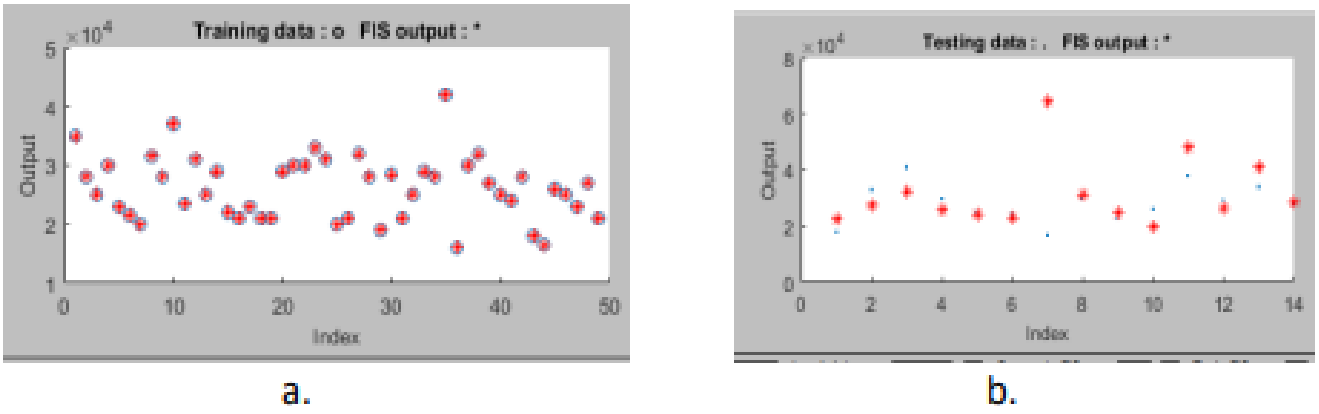
Table 4. Rules for rent estimation.

Rule no.	PC1	PC2	PC3	PC4	Estimated Rent(Around) (tk)
1.	SMALL	SMALL	SMALL	SMALL	38750
2.	SMALL	SMALL	SMALL	LARGE	8438
3.	SMALL	SMALL	LARGE	SMALL	28200
4.	SMALL	SMALL	LARGE	LARGE	24120
5.	SMALL	LARGE	SMALL	SMALL	6114
6.	SMALL	LARGE	SMALL	LARGE	18000
7.	SMALL	LARGE	LARGE	SMALL	54170
8.	SMALL	LARGE	LARGE	LARGE	15510
9.	LARGE	SMALL	SMALL	SMALL	4267
10.	LARGE	SMALL	SMALL	LARGE	62090
11.	LARGE	SMALL	LARGE	SMALL	101200
12.	LARGE	SMALL	LARGE	LARGE	25540
13.	LARGE	LARGE	SMALL	SMALL	71390

Result Analysis

We divide the 63 cases into two groups. One group contains 49 cases and another contains 14 cases. We perform PCA on both groups separately and get 4 PCs in both cases. We use the PCA outputs of 49 data

for training purpose, and the PCA outputs of 14 cases for the testing purpose. We use 5 membership functions and 40 epochs in training case.



a. b.

Fig 6. FIS output plotted against training data (a) and testing Data (b).

It can be seen from figure 6(a) that the plotting of the FIS output accurately matches with the plotting of the training data. In this case, the average training error is 0.50542. Though training data plotting matched with the FIS output with a small average error but from figure 6(b) it is seen that the average testing error is comparatively high. The reason behind this is the shortage of testing data. As we manage to collect only 63 data sets, the 77% of data has been used as training set and remaining 22% data is used as testing set. The data set of the14 flats hardly justify the range of the training data to begin with [2]. From the 14 testing data set we took some cases to compare their actual rent with the ANFIS generated rents. Table 5 depicts this

Table 5. Comparison of rents of random 5 test cases

Test Case No.	PC1	PC2	PC3	PC4	Source	Rent
1.	0.131	-1.24	-0.424	-0.664	ANFIS	28,200
					ACTUAL	28,000
2.	-0.7	0.105	0.795	0.759	ANFIS	25,100
					ACTUAL	25,000
3.	0.522	-1.66	-0.912	0.0732	ANFIS	27,600
					ACTUAL	30,000
4.	-0.211	0.227	1.77	-2.13	ANFIS	22,300
					ACTUAL	24,500
5.	-0.945	0.656	0.709	1.13	ANFIS	25,400
					ACTUAL	23,000

From Table 5, we see that case2 contains PC1, PC2, PC3 and PC4 values of - 0.700425108, 0.105194481, 0.795480446 and 0.758559801 respectively with rent of 25,000tk. Approximate values of the PC1, PC2, PC3 and PC4 were inputted in the rule view of the ANFIS. As it can be seen from Table 5, ANFIS outputs nearly the same value of 25,100tk. Therefore, from table 5 it can be seen that the ANFIS generated estimated rents are fair predictions of the actual rents. We conduct a surface graph shape test. We generate the surface view from PCA outputs from the ANFIS of 49 training data sets. Then we use the PCA outputs of the remaining 14 data sets to directly plot a surface graph using the software Graphing Calculator 3D separately. The axis ranges were kept same to compare the surface graphs on both cases. PC2, PC3, and PC4 is plotted along Y axis in figure 7, 8 and 9 respectively and along X-axis PC1 is plotted in all figures. Estimated Rent is in the Z- axis. It can be seen that there is a good similarity between the two graphs in shapes. The small data set (14 flats) generated surface looks like a small part of the large data (49 flats) set generated surface. The dissimilarities could have been decreased if the data sets could be made of the same size. This proves that the ANFIS generated results reflect similarity with the practical scenario.

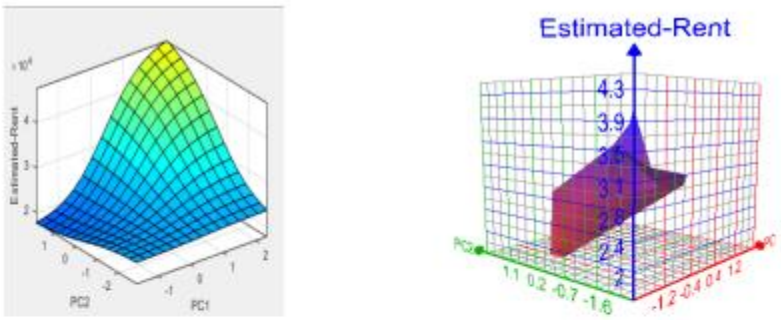


Fig 7. ANFIS generated smooth surface view (left) and Graphing Calculator 3D generated rough surface graph (right) for PC1 vs. PC2 vs. estimated rent

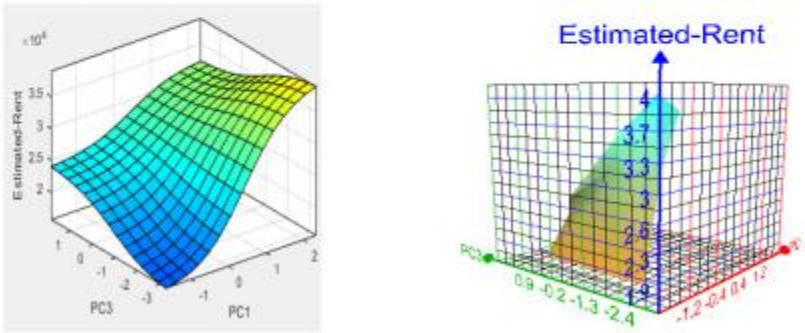


Fig 8. ANFIS generated smooth surface view (left) and Graphing Calculator 3D generated rough surface graph (right) for PC1 vs. PC3 vs. estimated rent

Conclusion: The creation of membership functions and fuzzy rules are one of the major challenges in implementation of fuzzy logic. In case of simple applications, it is feasible to construct these using common sense or domain knowledge, but in complicated system where there is huge volume of data set, it is difficult to predict. This paper represents the initial attempt to create a system to estimate flat rent based on Sugeno model. This research paper is focused on finding out estimated rent for flats. By using 3 types result analysis, it is shown that the process is factual and the results are nearly similar. This paper found accuracy while observing average training error, comparing ANFIS generated rent and actual rent.

Date: 24 / 11 / 2021 **Signature of faculty in-charge**