# Academic performance evaluation using soft computing techniques

**Ramjeet Singh Yadav[1,*], P. Ahmed[1], A. K. Soni[1] and Saurabh Pal[2]**

[1]Department of Computer Science and Engineering, SET, Sharda University, Greater Noida 201 306, India
[2]Department of MCA, Purvanchal University, Jaunpur 222 002, India

**This article presents a study of academic performance evaluation using soft computing techniques inspired by the successful application of K-means, fuzzy C-means (FCM), subtractive clustering (SC), hybrid subtractive clustering-fuzzy C-means (SC-FCM) and hybrid subtractive clustering-adaptive neuro fuzzy inference system (SC-ANFIS) methods for solving academic performance evaluation problems. Modelling of students' academic performance is a difficult optimization problem. We explore the applicability of K-means and FCM, SC, hybrid SC-FCM and SC-ANFIS clustering methods to the new student's allocation problem, which allocates new students into some classes that consist of similar students and the number of students in each class not exceeding its maximum capacity. The models were combined with fuzzy logic techniques to analyse the students' results. In this article, we have conducted clustering based computational experiments to analyse the effects of the different clustering algorithms like K-means, FCM, SC, hybrid SC-FCM and hybrid SC-ANFIS clustering methods for modelling students' academic performance evaluation. Based on the comparison of the results, it is found that the hybrid SC-ANFIS clustering is better than the other methods.**

**Keywords:** Academic performance evaluation, clustering algorithms, fuzzy logic, soft computing techniques.

THE student academic performance evaluation problem can be considered as a clustering problem where clusters (or classes) are formed on the basis of intelligence of students and predefined capacity class size. Intelligence-based grouping is essential for maintaining the homogeneity of the group; otherwise it would be difficult to provide good educational services to the highly diverse student population. Moreover, homogenous grouping of students having similar ranking (or some other measures) into classes would further make the academic performance results fairer, realistic and comparable. The existing practice of score aggregation-based students' similarity or their rank is least realistic because scores are assembled from different score combinations. Universities used grade point average (GPA), an example of score aggregation-based measure, as a major criterion for student selection. Most universities consider 3.0 and above GPA as an indicator of good academic performance. Hence, it remains the most common factor used by the academic planners to evaluate progression in an academic environment[1], despite its limitations in providing a comprehensive view of the state of students' performance evaluation and simultaneously discovering important details from their continuous performance assessments[2]. Further, average score may lead to wrong conclusion (especially, when details of data from which it is computed are not given).

It has been observed that there are factors, other than academic, which pose barriers to students attaining and maintaining high scores. Therefore, grouping or clustering students using cognitive as well as affective factors into different categories and then defining performance measure may be a realistic approach. For example, consider a scenario where two students score 50, 60, 70 and 70, 60, 50 in three tests respectively. The average mark obtained by each is 60. Can we conclude, from the average, that intelligence level of both the students is same? Of course not! The data indicate that one student is improving, whereas the other is deteriorating consistently, i.e. one student is learning consistently from his experience. The example illustrates that the student ranking or modelling academic performance evaluation method should be based on class homogeneity – a viewpoint supported by other researchers. Zukhri and Omar[3] have reported successful application of genetic algorithm for solving difficult optimization problems in new students' allocation problem. In addition to such computational issues, as mentioned before, imprecision and vagueness in the data collection process also affect evaluation of the performance indicators. Unfortunately, this aspect is ignored in practice because generally hard computing-based processes, procedures and techniques are used in performance evaluation. Observation shows that soft computing techniques are more powerful and better suited in providing feasible solutions to the problems that deal with uncertainties and vagueness. For instance, fuzzy logic handles imprecision and uncertainty in a natural manner by providing a human-oriented knowledge representation, but it is weak in self-learning and generalization of rules. A combination of fuzzy logic and fuzzy clustering algorithms is expected to

*For correspondence. (e-mail: ramjeetsinghy@gmail.com)

eliminate this weakness. Now, their power is being investigated.

Recently, Mankad *et al.*[4] have reported an evolving rule-based model for identification of multiple intelligence. Their genetic–fuzzy hybrid model identifies human intelligence. Sreenivasarao and Yohannes[5] have developed a model for improving academic performance evaluation of students based on data warehousing and data mining techniques that use soft computing intensively. Their analysis indicates that the group homogeneity improves students' academic performance, thereby enhancing education quality. An artificial neural network (ANN) model[6], along with computation also derives meaning from imprecise data, extracts patterns and detects trends. This ability has added new dimensions in comprehending the complex phenomenon that is buried in students' data, which otherwise might have gone unnoticed using hard computing techniques.

In practice, whether phenomena discovery or performance indicator computation, its accuracy depends on the data quality that in turn depends on the accuracy of data collection process and representation techniques. In order to address the data-related issues in education domain, the use of fuzzy sets in students' answer-sheets evaluation was suggested[7,8]. Using vague sets instead of fuzzy sets to represent the vague marks of each question was also suggested[9,10], where the evaluator can use vague values to indicate the degree of satisfaction for each question. In fuzzy sets the membership evaluation (characteristics function definition) is a major issue. In order to apply the fuzzy set effectively in educational domain, there have been several efforts in defining the effective membership. Bai and Chen[11] defined fuzzy membership functions for fuzzy rules, while Law[12] used fuzzy numbers; more information on academic performance evaluation is available in the literature[13–30]. These works indicate that fuzzy logic, neural network and fuzzy neural network have already been employed in student modelling systems, but nothing or very little has been mentioned about automatic generation of fuzzy membership function.

The present article describes various methods for automatic generation of membership function for student academic performance evaluation using K-means, fuzzy C-means (FCM), subtractive clustering (SC), hybrid subtractive clustering-fuzzy C-means (SC-FCM) and hybrid subtractive clustering-adaptive neuro fuzzy inference system (SC-ANFIS), which yields the homogeneous clusters (or classes) of students.

## Data cluster analysis techniques for academic performance evaluation

The clustering problem can be stated simply as follows: Given a finite set of data $X$, develop a grouping scheme for grouping the objects into classes. In classical cluster analysis, these classes are required to form a partition of $X$ such that the degree of association is strong for data within blocks of the partition and weak for data in different blocks. However, this requirement is too strong in practical applications, and it is thus desirable to replace it with a weaker requirement. When the requirement of a crisp partition of $X$ is replaced with a weaker requirement of a fuzzy partition or a fuzzy pseudo partition on $X$, we refer to the emerging problem area as fuzzy clustering. Fuzzy pseudo partitions are often called fuzzy $C$-partitions, where $C$ designates the number of fuzzy classes in the partition[31]. Finding grouping or trying to categorize the data for humans is not a simple task. This is why some methods in soft computing have been proposed to solve difficult optimization problems such as students' academic performance evaluation. The five methods (commonly known as data clustering techniques) and their performances determined by root mean square error (RMSE) are described in detail in Appendix 1. The outcome of these methods is given below.

## Results and discussion

The proposed methods (K-means, FCM, SC, hybrid SC-FCM and hybrid SC-ANFIS) allocate new students to homogenous groups of specified maximum capacity and analyse effects of such allocations on the academic performance of students. In these methods, the dataset used for training and testing is marks of 100 students who appeared in semester-1 (sem-1), semester-2 (sem-2) and semester-3 (sem-3), out of which 50 datasets have been used for training and rest 50 datasets for testing purpose (Tables 1 and 2).

The MATLAB software (used for modelling students' academic performance evaluation based on maximum value of marks that refers to the level of performance) based classification of the grades in this experiment is shown in Table 3. The marks obtained by each student who appeared in sem-1, sem-2 and sem-3 examinations have to be converted to the normalized values. Normalized value is referred to a range of (0, 1) which can be obtained by dividing the marks for each semester examination with the total marks. The normalized value will be the input value for evaluation. In addition, Table 3 also shows the marks and their associated original grade and level of achievement. Table 4 shows marks of 15 new students for testing the proposed models.

### *K-means method*

The datasets shown in Tables 1 and 2 have been divided into different clusters using K-means clustering method with the help of MATLAB software. The students have been classified in five groups (clusters) – very high, high, average, low and very low. K-means clustering method works on finding the cluster centres by trying to minimize objective function (eq. (5), Appendix 1). It alternates between updating the membership matrix and updating the cluster centres (eqs (7) and (8) respectively, Appendix 1) until no further improvement in the objective function is noticed. Since the algorithm initializes the cluster centres randomly, its performance is affected by initial cluster centres. After the cluster centres are determined, the evaluation data vectors are assigned to their respective clusters according to the distance between each vector and each of the cluster centres. An error measure is then calculated; the RMSE is used for this

**Table 1.** Student training dataset

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Final marks (statistical method) | Observed output | Grade |
|---|---|---|---|---|---|---|
| 1 | 0.05 | 0.37 | 0.18 | 0.200 | 0.25 | E |
| 2 | 0.10 | 0.23 | 10.6 | 0.163 | 0.25 | E |
| 3 | 0.15 | 0.13 | 0.06 | 0.113 | 0.25 | E |
| 4 | 0.40 | 0.13 | 0.20 | 0.243 | 0.25 | E |
| 5 | 0.25 | 0.31 | 0.14 | 0.233 | 0.25 | E |
| 6 | 0.15 | 0.10 | 0.26 | 0.170 | 0.25 | E |
| 7 | 0.10 | 0.13 | 0.30 | 0.177 | 0.25 | E |
| 8 | 0.10 | 0.17 | 0.08 | 0.117 | 0.25 | E |
| 9 | 0.25 | 0.23 | 0.04 | 0.173 | 0.25 | E |
| 10 | 0.05 | 0.17 | 0.12 | 0.113 | 0.25 | E |
| 11 | 0.12 | 0.32 | 0.34 | 0.260 | 0.45 | D |
| 12 | 0.25 | 0.33 | 0.30 | 0.293 | 0.45 | D |
| 13 | 0.30 | 0.30 | 0.34 | 0.313 | 0.45 | D |
| 14 | 0.40 | 0.20 | 0.38 | 0.327 | 0.45 | D |
| 15 | 0.50 | 0.40 | 0.30 | 0.400 | 0.45 | D |
| 16 | 0.65 | 0.17 | 0.38 | 0.400 | 0.45 | D |
| 17 | 0.50 | 0.26 | 0.38 | 0.380 | 0.45 | D |
| 18 | 0.55 | 0.35 | 0.38 | 0.427 | 0.45 | D |
| 19 | 0.50 | 0.40 | 0.40 | 0.433 | 0.45 | D |
| 20 | 0.45 | 0.51 | 0.36 | 0.440 | 0.45 | D |
| 21 | 0.40 | 0.60 | 0.44 | 0.480 | 0.55 | C |
| 22 | 0.35 | 0.60 | 0.48 | 0.477 | 0.55 | C |
| 23 | 0.32 | 0.50 | 0.65 | 0.490 | 0.55 | C |
| 24 | 0.55 | 0.60 | 0.48 | 0.543 | 0.55 | C |
| 25 | 0.30 | 0.70 | 0.54 | 0.513 | 0.55 | C |
| 26 | 0.45 | 0.47 | 0.60 | 0.507 | 0.55 | C |
| 27 | 0.40 | 0.40 | 0.64 | 0.480 | 0.55 | C |
| 28 | 0.35 | 0.50 | 0.58 | 0.477 | 0.55 | C |
| 29 | 0.35 | 0.63 | 0.58 | 0.520 | 0.55 | C |
| 30 | 0.25 | 0.47 | 0.72 | 0.480 | 0.55 | C |
| 31 | 0.40 | 0.67 | 0.64 | 0.570 | 0.75 | B |
| 32 | 0.35 | 0.61 | 0.76 | 0.573 | 0.75 | B |
| 33 | 0.60 | 0.70 | 0.54 | 0.613 | 0.75 | B |
| 34 | 0.50 | 0.60 | 0.66 | 0.587 | 0.75 | B |
| 35 | 0.80 | 0.73 | 0.62 | 0.717 | 0.75 | B |
| 36 | 0.55 | 0.75 | 0.76 | 0.687 | 0.75 | B |
| 37 | 0.75 | 0.57 | 0.84 | 0.720 | 0.75 | B |
| 38 | 0.50 | 0.87 | 0.72 | 0.697 | 0.75 | B |
| 39 | 0.70 | 0.47 | 0.86 | 0.677 | 0.75 | B |
| 40 | 0.85 | 0.57 | 0.76 | 0.727 | 0.75 | B |
| 41 | 0.70 | 0.82 | 0.76 | 0.760 | 1.00 | A |
| 42 | 0.80 | 0.87 | 0.74 | 0.803 | 1.00 | A |
| 43 | 0.85 | 0.90 | 0.80 | 0.850 | 1.00 | A |
| 44 | 0.75 | 0.83 | 0.84 | 0.806 | 1.00 | A |
| 45 | 0.85 | 0.87 | 0.88 | 0.867 | 1.00 | A |
| 46 | 0.90 | 0.67 | 0.96 | 0.843 | 1.00 | A |
| 47 | 0.95 | 0.87 | 0.90 | 0.907 | 1.00 | A |
| 48 | 0.95 | 0.97 | 0.98 | 0.967 | 1.00 | A |
| 49 | 0.90 | 0.93 | 0.94 | 0.923 | 1.00 | A |
| 50 | 1.00 | 0.83 | 0.98 | 0.937 | 1.00 | A |

**Table 2.** Student testing dataset

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Final marks (statistical method) | Observed output | Grade |
|---|---|---|---|---|---|---|
| 1 | 0.05 | 0.34 | 0.16 | 0.183 | 0.25 | E |
| 2 | 0.02 | 0.45 | 0.46 | 0.310 | 0.45 | D |
| 3 | 0.23 | 0.45 | 0.19 | 0.290 | 0.45 | D |
| 4 | 0.34 | 0.43 | 0.46 | 0.410 | 0.45 | D |
| 5 | 0.05 | 0.23 | 0.11 | 0.130 | 0.25 | E |
| 6 | 0.17 | 0.96 | 0.48 | 0.537 | 0.55 | C |
| 7 | 0.61 | 0.98 | 0.94 | 0.843 | 1.00 | A |
| 8 | 0.29 | 0.97 | 0.57 | 0.610 | 0.75 | B |
| 9 | 0.74 | 0.90 | 0.93 | 0.857 | 1.00 | A |
| 10 | 0.52 | 0.34 | 0.69 | 0.517 | 0.55 | C |
| 11 | 0.33 | 0.39 | 0.37 | 0.363 | 0.45 | D |
| 12 | 0.06 | 0.21 | 0.22 | 0.163 | 0.25 | E |
| 13 | 0.15 | 0.74 | 0.35 | 0.413 | 0.45 | D |
| 14 | 0.48 | 0.76 | 0.50 | 0.580 | 0.75 | B |
| 15 | 0.81 | 0.89 | 0.97 | 0.890 | 1.00 | A |
| 16 | 0.79 | 0.92 | 0.98 | 0.890 | 1.00 | A |
| 17 | 0.28 | 0.66 | 0.87 | 0.603 | 0.75 | B |
| 18 | 0.23 | 0.84 | 0.23 | 0.433 | 0.45 | D |
| 19 | 0.08 | 0.39 | 0.14 | 0.203 | 0.25 | E |
| 20 | 0.19 | 0.33 | 0.64 | 0.387 | 0.45 | D |
| 21 | 0.58 | 0.64 | 0.98 | 0.733 | 0.75 | B |
| 22 | 0.39 | 0.25 | 0.65 | 0.430 | 0.45 | D |
| 23 | 0.43 | 0.39 | 0.65 | 0.490 | 0.55 | C |
| 24 | 0.52 | 0.94 | 0.66 | 0.707 | 0.75 | B |
| 25 | 0.68 | 0.79 | 0.94 | 0.800 | 1.00 | A |
| 26 | 0.48 | 0.77 | 0.51 | 0.587 | 0.75 | B |
| 27 | 0.01 | 0.43 | 0.13 | 0.190 | 0.25 | E |
| 28 | 0.21 | 0.31 | 0.81 | 0.443 | 0.45 | D |
| 29 | 0.45 | 0.75 | 0.53 | 0.577 | 0.75 | B |
| 30 | 0.65 | 0.97 | 0.79 | 0.803 | 1.00 | A |
| 31 | 0.34 | 0.71 | 0.49 | 0.513 | 0.55 | C |
| 32 | 0.13 | 0.25 | 0.07 | 0.150 | 0.25 | E |
| 33 | 0.16 | 0.23 | 0.78 | 0.390 | 0.45 | D |
| 34 | 0.27 | 0.59 | 0.35 | 0.403 | 0.45 | D |
| 35 | 0.51 | 0.31 | 0.58 | 0.467 | 0.55 | C |
| 36 | 0.48 | 0.89 | 0.73 | 0.700 | 0.75 | B |
| 37 | 0.67 | 0.63 | 0.92 | 0.740 | 0.75 | B |
| 38 | 0.57 | 0.88 | 0.85 | 0.767 | 1.00 | A |
| 39 | 0.66 | 0.96 | 0.99 | 0.870 | 1.00 | A |
| 40 | 0.43 | 0.79 | 0.41 | 0.543 | 0.55 | C |
| 41 | 0.78 | 0.87 | 0.78 | 0.810 | 0.99 | A |
| 42 | 0.55 | 0.21 | 0.56 | 0.440 | 0.45 | D |
| 43 | 0.07 | 0.38 | 0.36 | 0.270 | 0.45 | D |
| 44 | 0.21 | 0.87 | 0.23 | 0.437 | 0.45 | D |
| 45 | 0.78 | 0.78 | 0.97 | 0.843 | 1.00 | A |
| 46 | 0.16 | 0.98 | 0.36 | 0.500 | 0.55 | C |
| 47 | 0.15 | 0.45 | 0.12 | 0.240 | 0.25 | E |
| 48 | 0.39 | 0.21 | 0.12 | 0.240 | 0.25 | E |
| 49 | 0.37 | 0.59 | 0.57 | 0.510 | 0.55 | C |
| 50 | 0.06 | 0.45 | 0.03 | 0.180 | 0.25 | E |

purpose (eq. (20), Appendix 1). The results of this method are given in Table 5 and the objective function values are shown in Figure 1.

It may be noted that three students belong to cluster-1, three students belong to cluster-2, five students belong to cluster-3, two students belong to cluster-4 and one student belongs to cluster-5 (Table 5). The drawback of K-means clustering method is that it cannot calculate the total marks of a student. Such a problem may be solved by the FCM, SC, hybrid SC-FCM and hybrid SC-ANFIS clustering methods.

**Table 3.** Marks and their associated original grade and level of achievement

| Sl no. | Marks | Grade | Level of achievement |
|---|---|---|---|
| 1 | 0.76–1.00 | A | Cluster-1 (very high) |
| 2 | 0.56–0.75 | B | Cluster-2 (high) |
| 3 | 0.46–0.55 | C | Cluster-3 (average) |
| 4 | 0.26–0.45 | D | Cluster-4 (low) |
| 5 | 0.00–0.25 | E | Cluster-5 (very low) |

**Table 4.** Dataset of students' score in sem-1, sem-2 and sem-3

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Final marks (statistical method) | Grade |
|---|---|---|---|---|---|
| 1 | 0.100 | 0.233 | 0.200 | 0.178 | E |
| 2 | 0.500 | 0.167 | 0.120 | 0.112 | E |
| 3 | 0.150 | 0.133 | 0.180 | 0.154 | E |
| 4 | 0.450 | 0.267 | 0.400 | 0.372 | D |
| 5 | 0.350 | 0.333 | 0.300 | 0.328 | D |
| 6 | 0.350 | 0.500 | 0.380 | 0.410 | D |
| 7 | 0.450 | 0.433 | 0.540 | 0.474 | C |
| 8 | 0.500 | 0.400 | 0.500 | 0.467 | C |
| 9 | 0.450 | 0.500 | 0.580 | 0.510 | C |
| 10 | 0.500 | 0.700 | 0.620 | 0.607 | B |
| 11 | 0.650 | 0.700 | 0.740 | 0.697 | B |
| 12 | 0.850 | 0.600 | 0.760 | 0.737 | B |
| 13 | 0.950 | 0.767 | 0.860 | 0.859 | A |
| 14 | 0.850 | 0.833 | 0.960 | 0.881 | A |
| 15 | 0.900 | 0.900 | 0.980 | 0.927 | A |

**Table 5.** Students' academic performance results using K-means method

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Grade based on K-means |
|---|---|---|---|---|
| 1 | 0.100 | 0.233 | 0.200 | D |
| 2 | 0.500 | 0.167 | 0.120 | E |
| 3 | 0.150 | 0.133 | 0.180 | D |
| 4 | 0.450 | 0.267 | 0.400 | C |
| 5 | 0.350 | 0.333 | 0.300 | C |
| 6 | 0.350 | 0.500 | 0.380 | C |
| 7 | 0.450 | 0.433 | 0.540 | C |
| 8 | 0.500 | 0.400 | 0.500 | C |
| 9 | 0.450 | 0.500 | 0.580 | C |
| 10 | 0.500 | 0.700 | 0.620 | B |
| 11 | 0.650 | 0.700 | 0.740 | B |
| 12 | 0.850 | 0.600 | 0.760 | B |
| 13 | 0.950 | 0.767 | 0.860 | A |
| 14 | 0.850 | 0.833 | 0.960 | A |
| 15 | 0.900 | 0.900 | 0.980 | A |

## *FCM method*

The baseline data (Tables 1 and 2) are divided into different clusters using FCM clustering using weighting exponent $m = 2$. The clustering number of the FCM method was initiated to 5, indicating availability of five rules. It consists of 15 instances, involving three conditional attributes: sem-1, sem-2 and sem-3, and five possible classification outcomes: clusters-1 to 5 (Table 6).

Noticeable is that the first student has been assigned performance index as 0.354 in FCM method (Table 6). Similarly, the fifth assigned performance index is 0.45. Figure 2 shows a plot of the objective function values. The objective function evolution suggests that the FCM method is better than K-means method. The FCM method provided faster convergence and higher accuracy for students' academic performance evaluation based on the following five rules.

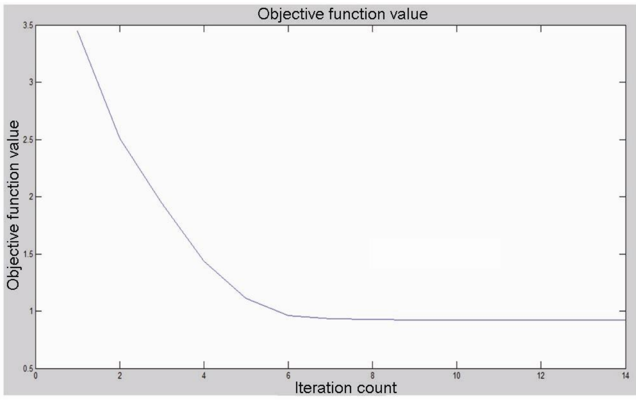


**Figure 1.** Objective function values of the K-means method.

**Table 6.** Students' academic performance results using FCM method

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Output | Grade |
|---|---|---|---|---|---|
| 1 | 0.100 | 0.233 | 0.200 | 0.354 | D |
| 2 | 0.500 | 0.167 | 0.120 | 0.358 | D |
| 3 | 0.150 | 0.133 | 0.180 | 0.357 | D |
| 4 | 0.450 | 0.267 | 0.400 | 0.457 | C |
| 5 | 0.350 | 0.333 | 0.300 | 0.449 | D |
| 6 | 0.350 | 0.500 | 0.380 | 0.500 | C |
| 7 | 0.450 | 0.433 | 0.540 | 0.555 | B |
| 8 | 0.500 | 0.400 | 0.500 | 0.517 | C |
| 9 | 0.450 | 0.500 | 0.580 | 0.608 | B |
| 10 | 0.500 | 0.700 | 0.620 | 0.687 | B |
| 11 | 0.650 | 0.700 | 0.740 | 0.765 | A |
| 12 | 0.850 | 0.600 | 0.760 | 0.788 | A |
| 13 | 0.950 | 0.767 | 0.860 | 0.877 | A |
| 14 | 0.850 | 0.833 | 0.960 | 0.866 | A |
| 15 | 0.900 | 0.900 | 0.980 | 0.871 | A |

(a) If sem-1, sem-2 and sem-3 are all in cluster-1, then academic performance is in cluster-1.

(b) If sem-1, sem-2 and sem-3 are all in cluster-2, then academic performance is in cluster-2.

(c) If sem-1, sem-2 and sem-3 are all in cluster-3, then academic performance is in cluster-3.

(d) If sem-1, sem-2 and sem-3 are all in cluster-4, then academic performance is in cluster-4.

(e) If sem-1, sem-2 and sem-3 are all in cluster-5, then academic performance is in cluster-5.

The first rule implies that the inputs to the FCM method; i.e. sem-1, sem-2 and sem-3, strongly belong to their respective cluster-1 membership functions and the student performance strongly belong to its cluster-1. The significance of the rule is that it succinctly maps cluster-1 in the input space to cluster-1 in the output space. Similarly, the second rules map cluster-2 in the input space to cluster-2 in the output space. If a datapoint closer to the first cluster, or in other words having strong membership to the first cluster, is fed as input to FCM, then rule 1 will operate predominantly than the rule 2. An input with strong membership to the second cluster will result in the operation of rule 2 predominantly than the other four rules. The outputs of the rules are then used to generate the output of the FCM method through the output membership functions. One output of the FCM, student performance, has five linear membership functions representing the five clusters identified by the FCM method. The coefficients of the linear membership functions though are not taken directly from the cluster centres. Instead, they are estimated from the dataset using least squares estimation technique in Tskagi–Sugeno (T–S) fuzzy model[32]. Consisting of a number of input–output linear regression models in each subspace, a T–S model can be built by means of fuzzy rule based on descriptions of input–output measurements of the academic performance evaluation. We conclude that the FCM method is an effective way to establish fuzzy inference rules described in the above-mentioned rules. However, due to multiple iterations and various eigen vectors, the FCM method suffers heavy computational burdens and is time-consuming. It is also highly sensitive to the initialization treatment, which usually requires a priori knowledge of the cluster numbers to form the initial cluster centres. Such limitations can be mitigated by the subtractive clustering based T–S fuzzy model[32] and hybrid SC-FCM method.

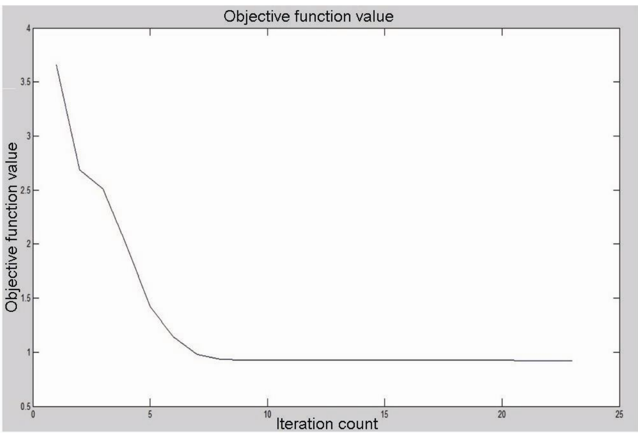


**Figure 2.** Objective function values of FCM.

**Table 7.** RMSE of training and testing datasets

| Training and testing RMSE | SC |
|---|---|
| Training | 0.039 |
| Testing | 0.107 |

## *Subtractive clustering method*

In SC method the baseline data (Tables 1 and 2) are divided into different clusters involving 100 instances, three conditional attributes: sem-1, sem-2 and sem-3, and five possible classification outcomes: clusters-1 to 5. For the sake of simplicity, only five linguistic labels, similar to the classification outcomes are used to represent student achievements. Clearly, the SC gives better fuzzification. Note that the given definition of the fuzzy sets is obtained solely on the basis of the normal distribution of the crisp marks. Table 7 shows the RMSE of training and testing datasets of the SC method.

The students' academic performance results using SC based on T–S fuzzy model[32] are given in Table 8. For example, it shows that if the first student has got 0.10 marks in sem-1, 0.23 marks in sem-2 and 0.20 marks in sem-3, then the performance of that student is 0.276 in the SC method. Similarly, the fifth student has a performance value of 0.415 in the SC method.

**Table 8.** Students' academic performance results using SC method

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Output | Grade |
|---|---|---|---|---|---|
| 1 | 0.100 | 0.233 | 0.200 | 0.276 | D |
| 2 | 0.500 | 0.167 | 0.120 | 0.219 | E |
| 3 | 0.150 | 0.133 | 0.180 | 0.253 | D |
| 4 | 0.450 | 0.267 | 0.400 | 0.479 | C |
| 5 | 0.350 | 0.333 | 0.300 | 0.415 | D |
| 6 | 0.350 | 0.500 | 0.380 | 0.503 | C |
| 7 | 0.450 | 0.433 | 0.540 | 0.550 | C |
| 8 | 0.500 | 0.400 | 0.500 | 0.544 | C |
| 9 | 0.450 | 0.500 | 0.580 | 0.553 | B |
| 10 | 0.500 | 0.700 | 0.620 | 0.767 | A |
| 11 | 0.650 | 0.700 | 0.740 | 0.768 | A |
| 12 | 0.850 | 0.600 | 0.760 | 0.817 | A |
| 13 | 0.950 | 0.767 | 0.860 | 0.943 | A |
| 14 | 0.850 | 0.833 | 0.960 | 1.080 | A |
| 15 | 0.900 | 0.900 | 0.980 | 1.070 | A |

**Table 9.** Students' academic performance based on FCM and SC-FCM methods

| | | | | FCM | | SC-FCM | |
|---|---|---|---|---|---|---|---|
| Sl no. | Sem-1 | Sem-2 | Sem-3 | Output | Grade | Output | Grade |
| 1 | 0.100 | 0.233 | 0.200 | 0.516 | C | 0.354 | D* |
| 2 | 0.500 | 0.167 | 0.120 | 0.518 | C | 0.469 | C |
| 3 | 0.150 | 0.133 | 0.180 | 0.517 | C | 0.357 | D* |
| 4 | 0.450 | 0.267 | 0.400 | 0.510 | C | 0.457 | C |
| 5 | 0.350 | 0.333 | 0.300 | 0.516 | C | 0.449 | D* |
| 6 | 0.350 | 0.500 | 0.380 | 0.524 | C | 0.500 | C |
| 7 | 0.450 | 0.433 | 0.540 | 0.571 | B | 0.556 | B |
| 8 | 0.500 | 0.400 | 0.500 | 0.511 | C | 0.517 | C |
| 9 | 0.450 | 0.500 | 0.580 | 0.613 | B | 0.609 | B |
| 10 | 0.500 | 0.700 | 0.620 | 0.688 | B | 0.686 | B |
| 11 | 0.650 | 0.700 | 0.740 | 0.720 | B | 0.765 | A* |
| 12 | 0.850 | 0.600 | 0.760 | 0.729 | B | 0.783 | A* |
| 13 | 0.950 | 0.767 | 0.860 | 0.710 | B | 0.876 | A* |
| 14 | 0.850 | 0.833 | 0.960 | 0.725 | B | 0.865 | A* |
| 15 | 0.900 | 0.900 | 0.980 | 0.720 | B | 0.870 | A* |

*Improve grade.

### *Hybrid subtractive clustering-fuzzy C-means method*

The baseline data (Tables 1 and 2) are divided into different clusters using hybrid SC-FCM method. This method has been trained by training data (Table 1) and tested by testing data (Table 2). It consists of 50 instances, involving three conditional attributes: sem-1, sem-2 and sem-3, and five possible classification outcomes: clusters-1 to 5. The primary assumption is that the partitions chosen by subtractive clustering are those best possible to represent the training data. Clearly, subtractive clustering has given better fuzzification. Note that the given definition of the fuzzy sets is obtained solely on the basis of the normal distribution of the crisp marks given. This ensures their comparison with other approaches.

In the present study, hybrid SC-FCM method deals with 50 datasets for training and 50 datasets for testing purpose, which are generated randomly within [0, 1] in two-dimensional space. The radius of hybrid SC-FCM method was specified as 0.5; the weighting exponent $m = 2$ and a termination criterion minimum improvement = 0.0000001. The hybrid SC-FCM method automatically generates appropriate clustering numbers according to the impact of each dimension of data on cluster centres, rather than demanding the number of clusters ahead. The clustering number of hybrid SC-FCM method was initiated to 5, which means five rules are available. On the contrary, inappropriate initial clustering number of the FCM method can lead to undesired results. To facilitate a fair comparison, the same dataset consisting of 15 instances and having the same features as the training dataset is used for both the methods. The five rules for the FCM method have also been generated by the hybrid SC-FCM method.

In the first rule, the inputs to the hybrid SC-FCM method, sem-1, sem-2 and sem-3, strongly belong to their respective cluster-1 and student performance (i.e. cluster-1). The significance of the rule is that it succinctly maps cluster-1 (very high) in the input space to cluster-1 in the output space. Similarly, the other four rules map cluster-2 (high), cluster-3 (average), cluster-4 (low) and cluster-5 (very low) in the input space to their respective clusters in the output space. If a datapoint is closer to the first cluster (or having strong membership to the first cluster), it will be fed as input to hybrid SC-FCM method; then rule (1) will operate predominantly than the other four rules. Similarly, an input with strong membership to the second cluster will result in the operation of second rule will with more firing strength than the other four rules, and so on. The outputs of the rules are then used to generate the output of the hybrid SC-FCM method through the output membership functions. One output of the hybrid SC-SCM method, student performance, has five linear membership functions representing the five clusters identified by subtractive clustering. The coefficients of the linear membership functions though are not taken directly from the cluster centres. Instead, they are estimated from the dataset using least squares estimation technique in T–S fuzzy model[32]. A comparison of FCM and SC-FCM methods in term of students' academic performance is shown in Table 9.

The first student has got performance index as 0.354 in hybrid SC-FCM method (Table 9). Similarly fifth student has got performance index 0.449. RMSE was employed to evaluate the accuracy of these models (for both training and testing data) which prevailed lower valued for hybrid SC-FCM, indicating its superiority.

The objective function evolution associated with the FCM and hybrid SC-FCM methods are shown in Figures 2 and 3, which indicates that hybrid SC-FCM method not only performs less iterations, but also achieves smaller value of the objective function. Thus the hybrid SC-FCM method provides faster convergence and higher accuracy for students' academic performance evaluation. Thus the proposed hybrid SC-FCM method provides better performance in comparison to FCM and other existing models for students' academic performance evaluation in the educational domain.

Figure 4 shows the model output and testing data by circles and lines respectively. It also shows that the model does not perform well on the testing data. Such limitations can be mitigated using the optimization capability of hybrid SC-ANFIS method to improve the model.

### *Hybrid SC-adaptive neuro-fuzzy inference system method*

To remove noise from both the basic data (Table 1, training data; and Table 2, testing data) model validation is needed to cross-validate the fuzzy inference system using testing dataset. The testing dataset is useful to check the generalization capability of the resulting T–S fuzzy model[32]. That is why the other 50 sets were used for testing after training was completed to verify the accuracy of the predicted values of academic performance evaluation. Marks obtained in sem-1, sem-2 and sem-3 are the inputs and the maximum values for the classification (see Tables 3 and 4) are the outputs of the system. Gaussian shapes are used for the membership function distribution for the

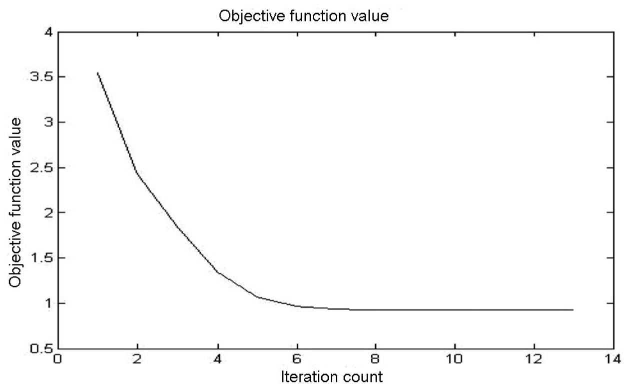


**Figure 3.** Objective function profiles of the hybrid SC-FCM method.

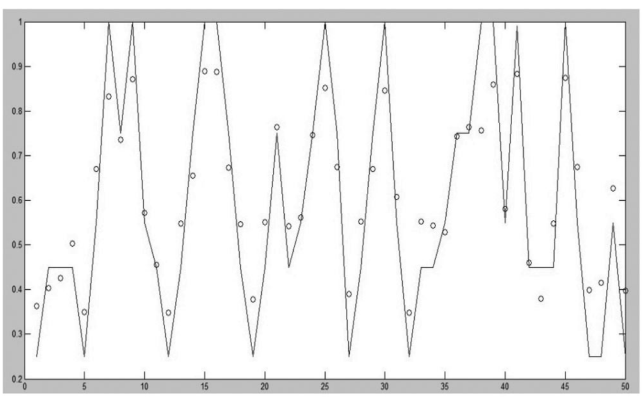

**Figure 4.** Hybrid SC-FCM output and testing data.

input variables. First-order T–S fuzzy model[32] is used in this study. The above three inputs of the fuzzy inference are classified into five fuzzy sets. Therefore, maximum number of fuzzy rules for this system can be five. During training in hybrid SC-ANFIS, 50 sets of experimental data were used to conduct 20 epochs of learning. ANFIS learning numbers for predicting academic performance are as follows: number of nodes: 46, number of linear parameters: 20, number of linear parameters: 30, total number of parameters: 50, number of training data pairs: 50, number of checking data pairs: 50, and number of fuzzy rules: 5. A hybrid SC-ANFIS based on first-order Sugeno fuzzy inference system is used to evaluate the students' academic performance in semester examinations. By employing the hybrid or back-propagation learning algorithm, hybrid SC-ANFIS can help obtain the optimal Gaussian membership functions.

The RMSE of SC-based FIS is 0.0392 for training data and 0.1092 for testing data, which indicates that the application of combined techniques of subtractive clustering and ANFIS achieved much satisfactory results in comparison to SC method for students' academic performance evaluation. The hybrid SC-ANFIS achieves slightly higher prediction accuracy than the SC method (Table 10).

Using a given input/output dataset, the hybrid SC-ANFIS constructs a T–S fuzzy model[32] whose membership function parameters are tuned (adjusted) using either a back-propagation algorithm alone or in combination with a least squares type of method. An important advantage of using a clustering method to find rules is that the resultant rules are more tailored to the input data. This reduces the problem of an excessive propagation of rules when the input data have large dimension. The RMSE value of testing and checking datasets of SC (0.1069) and hybrid SC-ANFIS (0.0874) shows that the RMSE of training and testing data sets is reduced against the SC

**Table 10.** Students' academic performance results using hybrid SC-ANFIS method

| Sl no. | Sem-1 | Sem-2 | Sem-3 | Output | Grade |
|---|---|---|---|---|---|
| 1 | 0.100 | 0.233 | 0.200 | 0.238 | E |
| 2 | 0.500 | 0.167 | 0.120 | 0.243 | E |
| 3 | 0.150 | 0.133 | 0.180 | 0.237 | E |
| 4 | 0.450 | 0.267 | 0.400 | 0.464 | C |
| 5 | 0.350 | 0.333 | 0.300 | 0.430 | D |
| 6 | 0.350 | 0.500 | 0.380 | 0.508 | C |
| 7 | 0.450 | 0.433 | 0.540 | 0.519 | C |
| 8 | 0.500 | 0.400 | 0.500 | 0.510 | C |
| 9 | 0.450 | 0.500 | 0.580 | 0.568 | B |
| 10 | 0.500 | 0.700 | 0.620 | 0.733 | B |
| 11 | 0.650 | 0.700 | 0.740 | 0.768 | A |
| 12 | 0.850 | 0.600 | 0.760 | 0.757 | A |
| 13 | 0.950 | 0.767 | 0.860 | 0.988 | A |
| 14 | 0.850 | 0.833 | 0.960 | 1.050 | A |
| 15 | 0.900 | 0.900 | 0.980 | 1.020 | A |

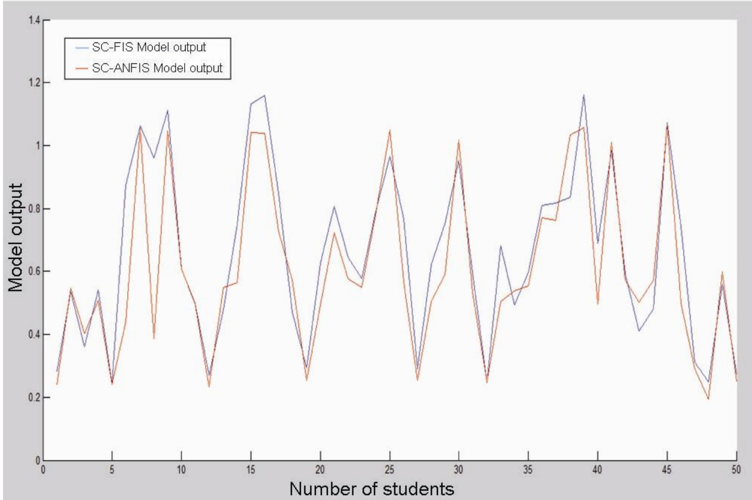


**Figure 5.** Comparison of output of SC and hybrid SC-ANFIS for testing datasets.

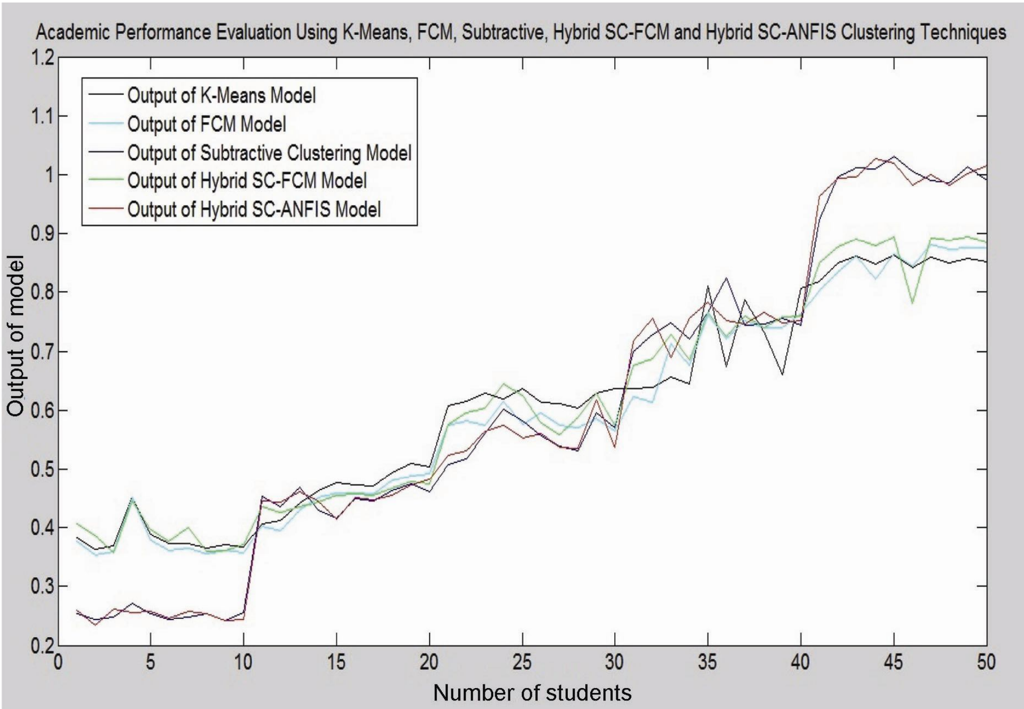


**Figure 6.** Comparison of K-means, FCM, SC, hybrid SC-FCM and hybrid SC-ANFIS.

method. Thus hybrid SC-ANFIS gives better results in comparison to SC method for academic performance evaluation. A comparison of output of SC and hybrid SC-ANFIS for testing dataset is shown in Figure 5.

### *Comparison of K-means, fuzzy C-means, SC, hybrid SC-FCM and hybrid SC-ANFIS clustering methods*

A comparison of all the methods shows that the first student belongs to cluster-4 (low) in K-means, FCM, SC, and hybrid SC-FCM methods, and to cluster-5 (very low) in hybrid SC-ANFIS method (Figure 6; Table 11). For the second student: cluster-5 (very low) in K-means, cluster-4 (low) in FCM, cluster-5 (very low) in SC, cluster-4 (low) in hybrid SC-FCM, and cluster-5 (very low) in hybrid SC-ANFIS methods respectively. This suggests that hybrid SC-ANFIS clustering method provides better results compared to other methods.

A summary of the five data clustering techniques and their results is given in Table 12 for academic performance evaluation. It shows that the RMSE hybrid SC-FCM is 0.0203 and 0.0874 for training and testing datasets respectively. These values are low in comparison to

**Table 11.** Comparison of K-means, fuzzy C-means, SC, hybrid SC-FCM and hybrid SC-ANFIS

| | | | | K-means | | FCM | | SC | | Hybrid SC-FCM | | Hybrid SC-ANFIS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sl no. | Sem-1 | Sem-2 | Sem-3 | Output | Grade | Output | Grade | Output | Grade | Output | Grade | Output | Grade |
| 1 | 0.100 | 0.233 | 0.200 | 0.321 | D | 0.354 | D | 0.276 | D | 0.345 | D | 0.238 | E* |
| 2 | 0.500 | 0.167 | 0.120 | 0.334 | E | 0.358 | D* | 0.219 | E* | 0.349 | D* | 0.243 | E* |
| 3 | 0.150 | 0.133 | 0.180 | 0.367 | D | 0.357 | D | 0.243 | E | 0.348 | D* | 0.237 | E* |
| 4 | 0.450 | 0.267 | 0.400 | 0.442 | C | 0.457 | C | 0.479 | C | 0.455 | C | 0.464 | C |
| 5 | 0.350 | 0.333 | 0.300 | 0.431 | C | 0.449 | D* | 0.415 | D* | 0.441 | D | 0.430 | D |
| 6 | 0.350 | 0.500 | 0.380 | 0.481 | C | 0.500 | C | 0.503 | C | 0.524 | C | 0.508 | C |
| 7 | 0.450 | 0.433 | 0.540 | 0.552 | C | 0.555 | B* | 0.550 | C* | 0.560 | C | 0.519 | D* |
| 8 | 0.500 | 0.400 | 0.500 | 0.503 | C | 0.517 | C | 0.544 | C | 0.508 | C | 0.510 | C |
| 9 | 0.450 | 0.500 | 0.580 | 0.571 | C | 0.608 | B* | 0.553 | B | 0.600 | B | 0.568 | B |
| 10 | 0.500 | 0.700 | 0.620 | 0.663 | B | 0.687 | B | 0.767 | A* | 0.678 | B* | 0.733 | B |
| 11 | 0.650 | 0.700 | 0.740 | 0.741 | B | 0.765 | A* | 0.768 | A | 0.788 | A | 0.768 | A |
| 12 | 0.850 | 0.600 | 0.760 | 0.754 | B | 0.788 | A* | 0.817 | A | 0.821 | A | 0.757 | A |
| 13 | 0.950 | 0.767 | 0.860 | 0.886 | A | 0.877 | A | 0.943 | A | 0.874 | A | 0.988 | A |
| 14 | 0.850 | 0.833 | 0.960 | 0.886 | A | 0.866 | A | 1.080 | A | 0.880 | A | 1.050 | A |
| 15 | 0.900 | 0.900 | 0.980 | 0.972 | A | 0.871 | A | 1.070 | A | 0.900 | A | 1.020 | A |

*Improve grade.

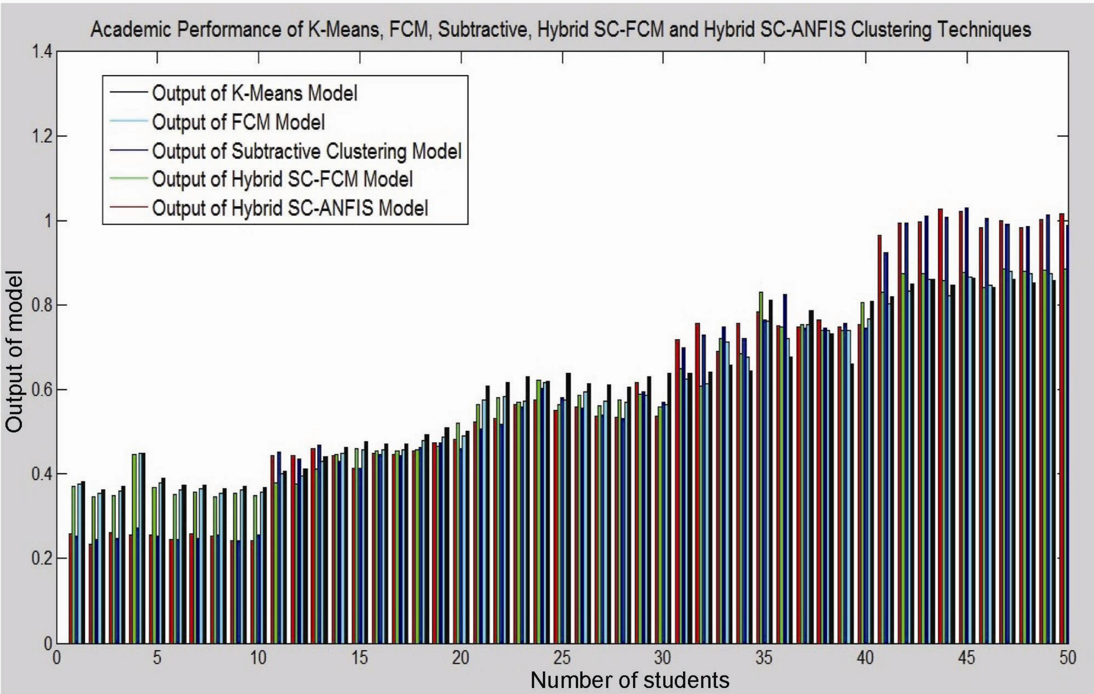


**Figure 7.** Bar chart of comparison of K-means, FCM, SC, hybrid SC-FCM and hybrid SC-ANFIS.

K-means, FCM, subtractive and hybrid SC-FCM clustering methods. Thus, it may be concluded that the hybrid SC-ANFIS gives better results for academic performance evaluation.

The following observations may be drawn from Table 12 for academic performance evaluation:

1. Hybrid SC-ANFIS clustering method shows higher accuracy and lower RMSE of training and testing datasets in comparison to the other four clustering techniques.
2. The FCM method gives results close to hybrid SC-FCM clustering method; yet hybrid SC-FCM method requires more computation time in comparison to FCM clustering method.
3. The subtractive clustering technique gives results closer to hybrid SC-ANFIS method; yet hybrid SC-ANFIS method requires more computation in comparison to SC method. Also, hybrid SC-ANFIS method gives better results compared to SC method.
4. The K-means clustering method gives poor results for training and testing datasets in comparison to the

**Table 12.** RMSE of K-means, FCM, SC-FCM, hybrid SC-FCM and hybrid SC-ANFIS methods

| Training and testing (RMSE) | K-means | FCM | SC-FCM | Hybrid SC-FCM | Hybrid SC-ANFIS |
|---|---|---|---|---|---|
| Training (RMSE) | 0.103 | 0.094 | 0.039 | 0.089 | 0.020 |
| Testing (RMSE) | 0.123 | 0.108 | 0.107 | 0.105 | 0.087 |

other four clustering techniques and RMSE is also high compared to other four clustering techniques.

The hybrid SC-ANFIS also automatically converts crisp data into fuzzy set and the model learned by ANN for further treatment of academic performance evaluation such as automatic calculation of membership function and automatic rule generation in development of dynamic fuzzy expert system for better evaluation of academic performance (Figure 7).

## Conclusion and future work

The present work provides qualitative methodology to compare the predictive power of clustering algorithm and the Euclidean distance using K-means, FCM, SC, hybrid SC-FCM and hybrid SC-ANFIS clustering methods for modelling academic performance evaluation. The hybrid SC-ANFIS is a more suitable technique in comparison to the other methods. It serves as a good benchmark to monitor the progress of students in educational modelling domain. It also improves decision-making ability of academic planners periodically by improving upon the academic results in the subsequent academic session. The proposed idea may be a starting point for the applicability of hybrid SC-ANFIS to analyse and model academic performance in the educational domain. The hybrid SC-ANFIS may serve as a potential tool for more effective and improved quality of education, better understanding of students' enrollment patterns in various courses, and amelioration of policies, and strategies for both students and teachers.

In future, the combination of hybrid SC-FCM and ANN (neuro-dynamic fuzzy expert system) may be used to evaluate the academic performance of both students and teachers in association with adaptive learning system and intelligent tutoring system for internet-based education and distance education.

## Appendix 1

### K-means clustering method

Based on iterative algorithm K-means clustering method involves moving clusters until the desired set is obtained[33] by classify data in a crisp sense. Define a family set $\{A_i,\ i = 1, 2, 3 \ldots, C\}$ as a partition of $X$, where the following set-theoretic forms can be applied for the partitions

$$\bigcup_{i=1}^{C} A_i = X, \quad (1)$$

$$A_i \cap A_j = \phi \qquad \forall i \neq j, \quad (2)$$

$$\phi \subset A_i \subset X \qquad \forall i, \quad (3)$$

where $X = \{x_1, x_2, x_3, \ldots, x_n\}$, a finite set space is comprised of the universe of data samples, and $C$ is the number of clusters to which classification has to be made. Obviously it may be noted

$$2 \leq C < n, \quad (4)$$

where $C = n$ classes just place each data sample into its own class. The objective function (or classification criteria) $J(U, v)$ is given as

$$J(U,v) = \sum_{k=1}^{n}\sum_{i=1}^{C} \chi_{ik}(d_{ik})^2, \quad (5)$$

where $U$ is the partition matrix, $v$ a vector of cluster centre and $d_{ik}$ a Euclidean distance measure between the $k$th data sample $x_k$ and $i$th cluster centre $v_i$, given by

$$d_{ik} = d(x_k - v_i) = \| x_k - v_i \| = \left[\sum_{j=1}^{m}(x_{kj} - v_{ij})\right]^{1/2}. \quad (6)$$

The algorithm is as below:

Step-I: Start with some initial configuration of prototypes $v_i$, $i = 1, 2, 3, \ldots, C$ (e.g. choose them randomly).

Step-II: Compute the value for $d_{ik}$ or the distance from the sample $x_k$ (a dataset) to the centre $c_i$, of the $i$th class, using eq. (4).

Step-III: Construct a partition matrix by assigning numeric values to $U$ according to the following rule

$$\chi_{ik} = \begin{cases} 1, & \text{if } d(x_k, v_i) = \min_{j \neq i} d(x_k, v_i), \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Step-IV: Update the prototype by computing the weighted average, which involves the entries of the partition matrix.

$$v_i = \frac{\sum_{k=1}^{N} \chi_{ik} x_k}{\sum_{k=1}^{N} \chi_{ik}}, \quad (8)$$

until the convergence criterion is met.

## FCM clustering method

The FCM generalizes the hard C-means algorithm to allow a point to partially belong to multiple clusters. It produces a constrained soft partition[34]. The extended objective function, denoted as $J$, is

$$J(U,V) = \sum_{i=1}^{k} \sum_{x_k \in X} (\mu_{C_i}(x_k))^m \| x_k - v_i \|^2, \quad (9)$$

where $U$ is a fuzzy partition of the dataset $X$ formed by $C_1, C_2, \ldots, C_k$. The parameter $m$ is a weight that determines the degree to which partial members of a cluster affect the clustering result. The FCM tries to find a good partition by searching for prototypes $v_i$ that minimize the objective function $J_m$. The FCM algorithms also need to search for membership functions $\mu_{C_i}$ that minimize $J$. A constrained fuzzy partition $\{C_1, C_2, \ldots, C_k\}$ can be a local minimum of the objective function $J$, only if the following conditions are satisfied[34]

$$\mu_{C_i}(x) = \frac{1}{\sum_{j=1}^{k} \left( \frac{\| x - v_i \|^2}{\| x - v_j \|^2} \right)^{\frac{1}{m-1}}} \qquad 1 \le i \le k, \ x \in X, \quad (10)$$

$$v_i = \frac{\sum_{x \in X} (\mu_{C_i}(x))^m x}{\sum_{x \in X}^{n} (\mu_{C_i}(x))^m} \qquad 1 \le i \le k, \quad (11)$$

$$\sum_{i=1}^{C} \| v_i^{\text{Previous}} - v_i \| \le \varepsilon. \quad (12)$$

In this way FCM updates the prototypes and the membership function iteratively using eqs (10) and (11) until a convergence criterion is achieved. The algorithm is as follows:

FCM($X$, $C$, $m$, $\varepsilon$)
  $X$: An unlabelled dataset
  $C$: the number of clusters to be formed
  $m$: the parameter in the objective function
  $\varepsilon$: A threshold for the convergence criteria.
  Initialize prototype $V = \{v_1, v_2, \ldots, v_c\}$
  Repeat $V^{\text{Previous}} \leftarrow V$
  Compute membership function using eq. (10).
  Update the prototype, $v_i$ in $V$ using eq. (11).
  Until $\sum_{i=1}^{C} \| v_i^{\text{Previous}} - v_i \| \le \varepsilon$
  Until convergence criterion is met.

## Subtractive clustering method

Originally based on the mountain method[35], the subtractive clustering is a fast, one-pass algorithm for estimating the number and centres of clusters for a set of data. Let $x$ be the dataset formed by concatenating the input dataset $X$ and the output dataset $Y$ of the system. Also, assume that each dimension of the data is normalized and mean dataset $x$ is bounded by hypercube. Subtractive clustering treats each point as a potential cluster centre and uses the following equation

$$D_i = \sum_{j=1}^{n} \exp\left[ \frac{| x_i - x_j |^2}{\left( \frac{r_a}{2} \right)^2} \right], \quad (13)$$

where $r_a$ defines the neighbourhood radius for each cluster. |.| is Euclidean distance and $n$ is the number of sampling points of the dataset $x$. Using eq. (13), subtractive algorithm computes the potential for each point. The point with the highest potential, denoted by $D_{c1}$ is selected as the first cluster centre $x_{c1}$. Next, the potential of each data point $x_i$ is updated as follows

$$D_i = D_i - D_{c1} \exp\left[ \frac{| x_i - x_{c1} |^2}{\left( \frac{r_b}{2} \right)^2} \right], \quad (14)$$

where $r_b$ represents the radius of the neighbourhood with significant potential reduction. Normally, $r_b$ should be chosen to be higher than $r_a$ to avoid closely spaced clusters. The next centre is spaced point with the highest potential. This process continues till a stopping criterion is met.

## Hybrid SC-FCM

The hybrid SC-FCM algorithm is presented as follows[36]:

Step 1: Calculate the density of every data point using eq. (13), and the highest density of the point is chosen as $x_{c1}$.

Step 2: Set $n_c = 1$, consider the highest potential of datapoint as $D_{c1}$ with location as $x_{c1}$ for the first cluster centre.

Step 3: Update each point potential using eq. (14).

Step 4: If, max $D_i \geq \varepsilon D_{c1}$ is true, accept $x_{c1}$ is the next cluster, which continues till one gets all cluster centres for the data.

Step 5: If, max $D_i \leq \varepsilon D_{c1}$ is true, go to step 4, otherwise, check if the point provides a good trade-off between having a sufficient potential and being sufficiently far away from existing cluster centres. If this is the case, this point is selected as the next cluster centre.

Step 6: Calculate $\mu_{C_i}(x)$ using eq. (10).

Step 7: Calculate and update the fuzzy cluster centre using eq. (11).

Step 8: Compute the objective function using eq. (9).

Step 9: If, eq. (12) is satisfied, then stop; otherwise, $k = k + 1$, otherwise go to step 4.

## SC-ANFIS method

The proposed SC-ANFIS model for academic performance evaluation based on subtractive clustering and ANFIS has learning capability. ANFIS proposed by Jang[37] has been implemented in the framework of adaptive networks. The ANFIS architecture with two inputs ($X_1$ and $X_2$), two rules and one output ($f$), for the first-order Sugeno fuzzy model, where each input is assumed to have two associated membership functions (Figure A1).

For a first-order Sugeno fuzzy model, a typical rule set with two fuzzy if – then rules can be expressed as

Rule (1): If $X_1$ is $A_1$ and $X_2$ is $B_1$ then $f_1 = m_1X_1 + n_1X_2 + q_1$.

Rule (2): If $X_1$ is $A_2$ and $X_2$ is $B_2$ then $f_2 = m_2X_1 + n_2X_2 + q_2$,

where $m_1$, $n_1$, $q_1$ and $m_2$, $n_2$, $q_2$ are the parameters of the output function. ANFIS consists of five layers and the functions of these layers are given below.

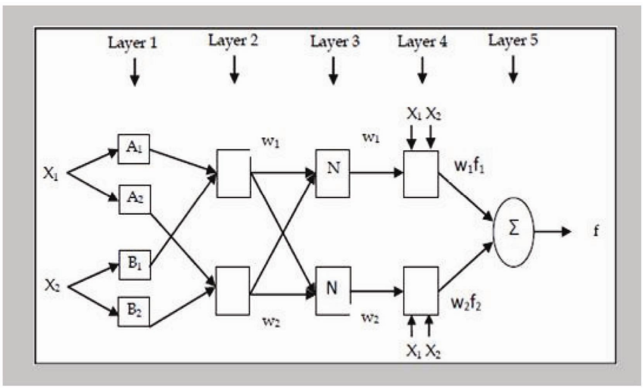


**Figure A1.** Structure of the proposed ANFIS model.

Layer 1: The node function of every node $i$ in this layer takes the form

$$O_i^1 = \mu_{A_i}(X), \tag{15}$$

where $X$ is the input to node $I$, $\mu_{A_i}(X)$ is the membership function (MF; which can be triangular, trapezoidal, Gaussian functions or other shapes) of the linguistic label $A_i$ associated with this node and $O_i$ is the degree of match to which the input $X$ satisfies the quantifier $A_i$. In the present study, the Gaussian shaped MFs defined below are utilized.

$$\mu_{A_i}(X) = \exp\left\{-\frac{1}{2}\frac{(X-c_i)^2}{\sigma_i^2}\right\}, \tag{16}$$

where $\{c_i, \sigma_i\}$ the parameters of the MFs governing the Gaussian are functions. The parameters in this layer are usually referred to as premise parameters.

Layer 2: Every node in this layer multiplies the incoming signals from layer 1 and sends the product out as follows

$$w_i = \mu_{A_i}(X_1) \times \mu_{B_i}(X_2), i = 1, 2, \tag{17}$$

where the output of this layer ($w_i$) represents the firing strength of a rule.

Layer 3: Every node $i$ determine the ratio of the $i$th rule's firing strength to the sum of all rules' firing strengths as

$$w = \frac{w_i}{w_1 + w_2}, \tag{18}$$

where the output of this layer represents the normalized firing strength.

Layer 4: Every node $i$ in this layer is an adaptive node with a node function of the form

$$o_i^4 = wf_i = w(m_iX_1 + n_iX_2 + q_i), \quad i = 1, 2, \tag{19}$$

where $w$ is the output to layer 3, and $\{m_i, n_i, q_i\}$ is the parameter set of this node. Parameters in this layer are referred to as consequent parameters.

Layer 5: There is only a single node in this layer that computes the overall output as the weighted average of all incoming signals from layer 4 as

$$O_i^5 = \sum_i wf_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, \quad i = 1. \tag{20}$$

## Calculation of RMSE

The performances of the above methods are determined by RMSE using the following equation

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(X_{\mathrm{obs}} - X_{\mathrm{model}})^2}{n}}, \quad (21)$$

where $X_{\mathrm{obs}}$ is the observed value and $X_{\mathrm{model}}$ is the modelled values.

1. Sansgiry, S. S., Bhosle, M. and Sail, K., Factors that affect academic performance among pharmacy students. *Am. J. Pharm. Educ.*, 2006, **70**(5), 231–243.
2. Oyelade, O. J., Oladipupo, O. O. and Obagbua, I. C., Application of K-means clustering algorithm for prediction of students' academic performance. *Int. J. Comput. Sci. Inf. Security*, 2010, **7**(1), 292–295.
3. Zukhri, Z. and Omar, K., Solving new student allocation problem with genetic algorithm: a hard problem for partition based approach. *Int. J. Soft Comput. Appl.*, 2008, **2**, 6–15.
4. Mankad, K., Sajja, P. S. and Akerkar, R., Evolving rules using genetic fuzzy approach: an educational case study. *Int. J. Soft Comput.*, 2011, **2**(1), 35–46.
5. Sreenivasarao, V. and Yohannes, G., Improving academic performance of students of Defense University based on data warehousing and data mining. *Global J. Comput. Sci. Technol.*, 2012, **12**(2), 201–209.
6. Afoayan, O. and El-Shamir Absalom, E., Design and implementation of student's information system for tertiary institutions using neural network techniques. *Int. J. Green Comput.*, 2010, **1**(1), 1–15.
7. Biswas, R., An application of fuzzy sets in students' evaluation. *Fuzzy Sets Syst.*, 1995, **74**(2), 187–194.
8. Zadeh, L. A., Fuzzy sets. *Inf. Control*, 1965, **8**, 338–354.
9. Gau, W. L. and Buehrer, D. J., Vague sets. *IEEE Trans. System, Man Cybernetics*, 1993, **23**(2), 610–614.
10. Wang, H. Y. and Chen, S. M., New methods for evaluating students' answerscripts using fuzzy numbers associated with degrees of confidence. In 2006 IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, 2006, pp. 1004–1009.
11. Bai, S. M. and Chen, S. M., Evaluating students' learning achievement using fuzzy membership functions and fuzzy rules. *Expert Syst. Appl.*, 2008, **34**, 399–410.
12. Law, C. K., Using fuzzy numbers in educational grading system. *Fuzzy Sets Syst.*, 1996, **83**, 311–323.
13. White, H., Learning in artificial neural networks: a statistical perspective. *Neural Comput.*, 1989, **1**, 425–464.
14. Chen, S. M. and Lee, C. H., New methods for students' evaluation using fuzzy sets. *Fuzzy Sets Syst.*, 1999, **104**, 209–218.
15. Stathacopoulou, R., Magoulas, G. D., Grigoriadou, M. and Samarakou, M., Neuro-fuzzy knowledge processing in intelligent learning environments for improved student diagnosis. *Inf. Sci.*, 2005, **170**(2–4), 273–307.
16. Chaudhari, O. K., Khot, P. G. and Deshmukh, K. C., Soft computing model for academic performance of teachers using fuzzy logic. *Br. J. Appl. Sci. Technol.*, 2012, **2**(2), 213–226.
17. Neogi, A., Mondal, A. C. and Mandal, S., A cascaded fuzzy inference system for university non-teaching staff performance appraisal. *J. Inf. Process. Syst.*, 2011, **7**(4), 595–612.
18. Yadav, R. S. and Singh, V. P., Modeling academic performance evaluation using soft computing techniques: a fuzzy logic approach. *Int. J. Comput. Sci. Eng.*, 2011, **3**(2), 676–686.
19. Upadhyay, M. S., Fuzzy logic based on performance of students in college. *J. Comput. Appl.*, 2012, **5**(1), 6–9.
20. Guh, Y. Y., Yang, M. S., Po, R. W. and Lee, E. S., Establishing performance evaluation structures by fuzzy relation based cluster analysis. *Comput. Math. Appl.*, 2008, **56**, 572–582.
21. Pavani, S., Gangadhar, P. V. S. S. and Gulhare, K. K., Evaluation of teacher's performance evaluation using fuzzy logic techniques. *Int. J. Comput. Trends Technol.*, 2012, **3**(2), 200–205.
22. Gokmen, E., Akinci, T. C., Tektas, M., Onat, N., Kocyigit, G. and Tektas, N., Evaluation of student performance in laboratory applications using fuzzy logic. *Proc. Soc. Behav. Sci.*, 2010, **2**, 902–909.
23. Hameed, I. A., Using Gaussian membership functions for improving the reliability and robustness of students' evaluation system. *Expert Syst. Appl.*, 2011, **38**(6), 7135–7142.
24. Baylari, A. and Montazer, G. A., Design a personalized e-learning system based on item response theory and artificial neural network approach. *Expert Syst. Appl.*, 2009, **36**, 8013–8021.
25. Cios, K. J., Pedrycz, W., Swiniarski, R. W. and Kurgan, L., *Data Mining: A Knowledge Discovery Approach*, Springer, 2007, pp. 263–265.
26. Posey, C. L. and Hawkes, L. W., Neural networks applied in the student model. *Intell. Syst.*, 1996, **88**, 275–298.
27. Gupta, C. R. and Dhawan, A. K., Diagnosis, modeling and prognosis of learning system using fuzzy logic and intelligent decision vectors. *Int. J. Comput. Appl.*, 2012, **37**(6), 975–987.
28. Stathacopoulou, R., Grigoriadou, M., Samarakou, M. and Mitoropoulou, D., Monitoring students' action and using teachers' expertise in implementing and evaluating the neural network-based fuzzy diagnostic model. *Expert Syst. Appl.*, 2007, **32**, 955–975.
29. Bhatt, R. and Bhatt, D., Fuzzy logic based student performance evaluation model for practical components of engineering institutions subjects. *Int. J. Technol. Eng. Educ.*, 2011, **8**(1), 1–7.
30. Ma, J. and Zhou, D., Fuzzy set approach to the assessment of student centered learning. *IEEE Trans. Educ.*, 2000, **43**(2), 112–120.
31. Gagula-Palalic, S. and Can, M., Fuzzy clustering models and algorithms for pattern recognition. Master's thesis, International University of Sarajevo, 2008, pp. 13–17.
32. Huaguang, Z., Jianhong, L. and Laijiu, C., The theory and application research on predictive generalized control. *Acta Autom. Sin.*, 1993, **19**(3), 9–17.
33. Hartigan, J. A. and Wong, M. A., A K-means clustering algorithm. *Appl. Stat.*, 1979, **28**, 100–108.
34. Bezdek, J. C., *Pattern Recognition and Fuzzy Objective Function Algorithms*, Plenum Press, 1981, pp. 65–86.
35. Chiu, S., Fuzzy model identification based on cluster estimation. *J. Intell. Fuzzy Syst.*, 1994, **2**(3), 267–278.
36. Hossen, J., Rahman, A., Sayeed, S., Samsuddin, K. and Rokhani, F., A modified hybrid fuzzy clustering algorithm for data partitions. *Austr. J. Basic Appl. Sci.*, 2011, **5**(8), 674–681.
37. Jang, J. S. R., ANFIS: adaptive-network based fuzzy inference system. *IEEE Trans. Syst., Man, Cybernetics*, 1993, **23**(3), 665–685.

ACKNOWLEDGEMENTS. We thank Dr Kiran K. Ravulakollu and Prof. N.B. (RTDC), Sharda University, Grater Noida, Prof. Lalmani and Dr C. P. Kushwaha (Ashoka Institute of Technology and Management, Varanasi) for help.

Received 2 April 2013; revised accepted 28 March 2014