Fuzzy Clustering and its Application in Education of Gifted Students

Introduction

Care of gifted students is currently a very relevant topic which affects most of the scientific disciplines. The aim of this article is to outline one of the possibilities of education and also the identification of gifted students using so-called fuzzy clustering. The aim of the fuzzy clustering is to create equal groups of students according to the type and level of their giftedness and consequent raising of effectiveness of the educational process. At the same time, the identification of student's giftedness is done.

In our article, the term "gifted student" is used for an individual whose results of education in the intellectual area exceed the average results of common population of individuals of a similar age from the identical educational environment. Our definition of the term "gifted student" is backed by the multidimensional definitions (multiple criteria of performance are observed during the identification of giftedness) based on manifested performance (only the openly manifested performance is observed).

In today's educational system, the integrative educational direction, which requires equal conditions in education of all pupils and students and the effort for the maximum inclusion of an individual into the mainstream education, prevails. Fuzzy clustering in education of gifted students may currently appear as practically unusable. However, fuzzy clustering may have a significant place in the integrative education as one of the possibilities and forms of individual educational approach to gifted students. The main purpose of fuzzy clustering is the occasional creation of work groups of students whose level and type of intellectual giftedness would be similar and who would be offered and adequate educational approach.

Related works and suggestion of solution of problem by cluster analysis

One possibility how to apply fuzzy clustering of the students according to their type and level of giftedness in integrative education is the use of B. S. Bloom's sophisticated individualizing system Mastery Learning (Guskey, 2012). Mastery Learning is based on the assumption that a certain standard of education may be achieved by all the students, given the appropriate conditions for learning, which is their grouping according to their type and level of giftedness and the subsequent application of different types of tasks based on the tasks' difficulty (Morgan, 2011). If we group the students according to their type and level of giftedness, the students' learning effectivity and motivation in the individual groups is maximally raised (Heward, 2013). Mastery Learning maintains that students must achieve a level of mastery in prerequisite knowledge before moving forward to learn subsequent information. If a student does not achieve mastery on the test, they are given additional support in learning and reviewing the information, and then tested again. Mastery learning methods suggest that the focus of instruction should be the time required for different students to learn the same material and achieve the same level of mastery. This is very much in contrast with classic models of teaching, which focus more on differences in students' ability and where all students are given approximately the same amount of time to learn and the same set of instructions.

The organization of the system is planned in such a way, that after the common education, the students are given feedback about how and in which areas of education they achieved the pedagogical goal. It is, for example, the evaluation methods such as the instructions for the didactic tests or structured observation of student's performance by a teacher etc. Based on the successfulness of the evaluation method and the type of giftedness, the students are subsequently divided into 3 or more subgroups where they work on individualized tasks (Figure 1). After that, all the groups are joined into the basic group, so that they can continue in following common tasks (Gentile, 2003).
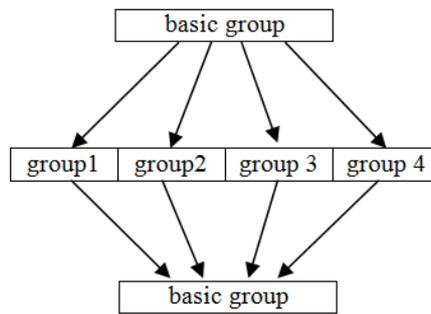
Fig. 1 Model of grouping according to the Mastery learning

During the process of evaluation students for the purpose to divide them into groups, several methods are used. Each evaluation method follows several criteria. For example didactic test could contain partial criteria as a learning quality, logical thinking, text comprehension etc. (Callahan, Renzulli, 2012).

During evaluation of partial criteria the broad model is suggested (Renzulli and Reis 2004). In the broad model the individual, who comes into the best group (group of gifted students), must fulfil all or the most of evaluation criteria, so the methodological problem is how to combine these each results.

In praxis and theory (Callahan and Renzulli 2012) is application of the additive model registered. In this model the partial outcomes are easily added for each individual in evaluation process. These outputs in a form of some total points are compared, so the individuals with the best results are nominated into the special education form. The advantage of additive model is quite easy evaluation, where we add each point together. On the other hand we add the criteria with different conditions and relevance (for example logical thinking and motivation), moreover the results from each different criteria could be inappropriately compensate and give mistaken results of identification process.
Hunsaker (Hunsaker 2012) notes, that different identification criteria cannot be added linearity. He suggests addition of selected criteria which plays key role for identification of gifted students and other less important criteria which are used tentatively.

To eliminate these disadvantages we suggest combining each result by using the cluster analysis. One of the main advantages of fuzzy clustering is the ability to express ambiguity in the assignment of objects to clusters. It generates fuzzy partitions of the data instead of hard partitions. Experimental results prove that fuzzy clustering seems also to be more robust in terms of local minima of the objective function. The advantages of fuzzy clustering model against added model are in the use of vague inputs and outputs terms and evaluation of non - linear dependence.

We found that there were no application of the fuzzy clustering during identification evaluation process of master learning according to analyze of available article database (EBSCO, XERXES and Proquest).

Cluster Analysis

A cluster analysis or a clustering is the task of grouping a set of objects in such a way that objects in the same group (called cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). Popular notions of clusters include groups with low distances among the cluster members. The fuzzy clustering could be also used in education of gifted students.

In hard clustering, data is divided into distinct clusters, where each data element belongs to exactly one cluster. In fuzzy clustering (also referred to as soft clustering), data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a

process of assigning these membership levels, and then using them to assign data elements to one or more clusters.

One of the most widely used fuzzy clustering algorithms is the Fuzzy C-Means algorithm (Bezdek, 1981). The Fuzzy C-Means algorithm attempts to partition a finite collection of n elements $X=\{x_1, x_2, ... , x_n\}$ into a collection of c fuzzy clusters with respect to some given criterion. Given a finite set of data, the algorithm returns a list of c cluster centers where each element and a partition matrix $W = w_{i,j} \in [0,1]$, $i=1,2, ... , n$, $j = 1, 2, ... , c$, where each element $w_{i,j}$ tells the degree to which element x_i belongs to cluster c_j. the fuzzy C-means aims to minimize an objective function. The standard function is

$$w_k = \frac{1}{\sum_j \frac{d(center_{k,x})^{2/(m-1)}}{d(center_{k,x})}} \qquad (1)$$

this differs from the k-means objective function by the addition of the membership values u_{ij} and the fuzzifier m. The fuzzifier m determines the level of cluster fuzziness. A large m results in smaller memberships w_{ij} and hence, fuzzier clusters. In the limit $m = 1$, the memberships w_{ij} converge to 0 or 1, which implies a crisp partitioning. In the absence of experimentation or domain knowledge, m is commonly set to 2. The basic Fuzzy C-means algorithm, given n data points $(x_1, x_2, ... , x_n)$, to be clustered, a number of c clusters with $(c_1, c_2, ... , c_n)$ and m the level of cluster fuzziness with.

In fuzzy clustering, every point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to just one cluster. Thus, points on the edge of a cluster may be in the cluster to a lesser degree than points in the center of cluster.

Any point x has a set of coefficients giving the degree of being in the k-th cluster $w_k(x)$. With fuzzy C-means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster

$$c_k = \frac{\sum_x w_k(x)x}{\sum_x w_k(x)} \qquad (2)$$

The degree of belonging, $w_k(x)$, is related inversely to the distance from x to the cluster centre as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest centre.

Case study
The application of the fuzzy logic model is realized on the case study concerning of education of gifted students. The solved problem is based on sorting of pupils according to the abilities. In other words, we have to find the very, middle and low gifted pupils. The data of the case study - inputs and outputs are following. The inputs are characterized by product's parameters such as *Speed and learning quality (S)*, *Logical thinking (T)* and *Learning quality (L)*. Input data are represented by 74 objects. See Table 1.

Table 1: Cluster data

Order	S	T	L	Order	S	T	L
1	3	3	3	38	3	3	3
2	2	2	2	39	2	2	3
3	2	2	3	40	4	4	3
4	2	2	2	41	4	4	4
5	4	4	4	42	3	3	3
6	2	2	2	43	2	2	2
7	3	2	3	44	2	2	3
8	4	4	4	45	2	1	1

9	2	2	2	46	3	3	3
10	3	3	4	47	2	2	2
11	2	2	1	48	4	4	3
12	2	2	2	49	4	5	4
13	4	4	4	50	3	3	1
14	2	2	3	51	1	2	2
15	3	2	4	52	2	3	3
16	2	3	2	53	1	2	3
17	2	2	3	54	4	4	3
18	5	5	3	55	1	1	1
19	4	3	3	56	4	3	4
20	3	2	3	57	4	4	3
21	3	3	4	58	2	2	2
22	3	2	3	59	3	3	4
23	4	4	3	60	1	2	1
24	5	5	5	61	4	4	4
25	3	3	3	62	2	2	2
26	4	4	4	63	3	3	4
27	3	4	3	64	2	2	2
28	3	3	4	65	2	2	3
29	4	3	3	66	5	5	3
30	2	2	2	67	4	4	3
31	3	3	4	68	3	3	3
32	5	5	5	69	3	2	3
33	2	2	2	70	2	2	2
34	4	3	3	71	3	3	2
35	1	2	2	72	5	5	5
36	2	2	2	73	3	3	3
37	4	4	3	74	3	3	3

The output will be the classification of gifted students according their characteristic to clusters. The software MATLAB and its Fuzzy Logic Toolbox was used for the software applications.

The number of clusters was set up to 4. The calculation is presented by graph of dependence of fitness function on number of iteration process and it proves to good optimization process. See Figure 2. The results are presented by coordinates of clusters and assignment of pupils to clusters according its *S*, *T* and *L*. See Figure 3.
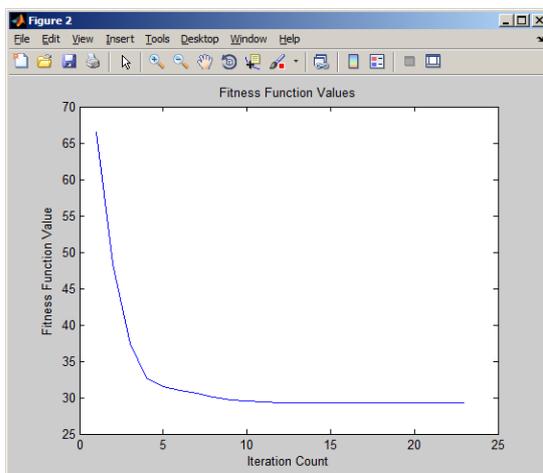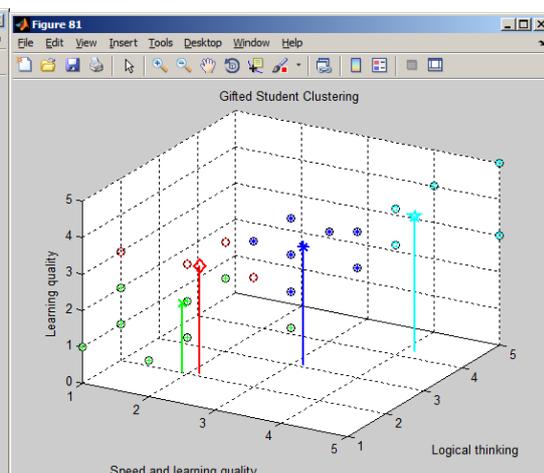

Fig. 2 Fitness function values

Fig. 3 Graph – gifted students clusters

When the calculation is finished the output results, the iteration count, the coordinates of centroids and assign of product to centroids are displayed. See Table 2.

Table 2: Cluster data

c1 = ×			c2 = ♦			c3 = ✳			c4 = ♣		
2	2	2	2	2	3	3	3	3	4	4	4
2	2	2	3	2	3	3	3	4	4	4	4
2	2	2	2	2	3	3	2	4	4	4	4
2	2	2	2	2	3	4	3	3	5	5	3
2	2	1	3	2	3	3	3	4	4	4	3
2	2	2	3	2	3	3	3	3	5	5	5
2	3	2	2	2	3	3	4	3	4	4	4
2	2	2	2	2	3	3	3	4	5	5	5
2	2	2	2	3	3	4	3	3	4	4	3
1	2	2	1	2	3	3	3	4	4	4	3
2	2	2	2	2	3	4	3	3	4	4	4
2	2	2	3	2	3	3	3	3	4	4	3
2	1	1				3	3	3	4	5	4
2	2	2				3	3	3	4	4	3
3	3	1				4	3	4	4	4	3
1	2	2				3	3	4	4	4	4
1	1	1				3	3	4	5	5	3
2	2	2				3	3	3	4	4	3
1	2	1				3	3	2	5	5	5
2	2	2				3	3	3			
2	2	2									
2	2	2									

Conclusion

The results are presented by centroids of four clusters marked ×,✳,♦,♣ and assignment of pupils to the clusters. The results presents the case where the cluster ✕ includes the group of pupils with *low S, low T* and *low L*, the cluster ✳ includes the group of pupils with *low S, low T* and *medium L*, the cluster ♦ includes the group of pupils with *medium S, medium T* and *medium L* and the cluster ♣ includes the group of pupils with *high S, high T* and *high L*. The fuzzy model enables to identify gifted pupils according their *S, T, L*. The tasks from practice lead to multi-dimensional ones, where their graphical presentation is impossible. It can be mentioned other inputs such as: broad vocabulary application, text comprehension, knowledge of subject language and literature, reading, speaking, activity and motivation.

The benefit of Fuzzy C-Means algorithm resides in the fact that it enables to process vague and imprecision data that are commonly used in pedagogical field.

The example mentioned above is an application of the use of fuzzy logic for identification of gifted people. The identification of gifted people plays an important role in educational processes.

References

Arfi, B. Linguistic Fuzzy Logic methods in Social Sciences. Studies in Fuzziness and Soft Computing, Volume 253, Springer, 185p., ISBN 978-3-642-42237-9.

Bezdek, J.C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. USA: Springer.

Dostál, P. (2011). Advanced Decision Making in Business and Public Services. Czech Republic: CERM Academic Publishing House.

Dostál, P. (2014). The Use of Soft Computing in Management. In Handbook of Research on Novel Soft Computing Intelligent Algorithms: Theory and Practical Applications USA: IGI Globe, 2013.

Guskey, T.R. ed. (2012). Benjamin S. Bloom – Portraits of an Educator. UK: Rowman and Littlefield Publishers.

Gentile, J.R. (2003). Standards an Mastery Learning. USA: Sage.

Heward, W.L. (2013). Exceptional Children. An Introduction to Special Education. Ohio: Pearson Education.

Morgan, K. (2011). Mastery Learning in the Science Classroom. USA: NSTA press.

Porter, L. (1999). Gifted Young Children. A guide for teachers and parents. Buckingham: Open University Press.

Hunsaker, S.L. (2012). The Theory and Practise of Identifying Students for Gifted and Talented Education Services. USA: Prufrock.

Zadeh, L.A. (1965). Fuzzy sets. Information and Control 8, 338-353.

Zadeh, L.A. (2012). A Definition of Soft Computing - adapted from L.A. Zadeh. Retrieved January 10, 2013, from http://www.soft-computing.de/def.html.