Evaluation of training process by Fuzzy techniques

Sorin BURLACU
The Bucharest Academy of Economic Studies, Romania

Abstract: *Through this paper we propose a new approach to teaching evaluation. In this research I started from the premise that education is a process of evaluating and interpreting from data source. The evaluation shows the level of understanding where reached at a given time.*

Relativity units of measurement and limits of current instruments that commensurate level of knowledge, led us to seek new methods and techniques to quantify the degree of knowledge of a phenomenon or object of knowledge.

Techniques that derived from fuzzy set theory have proved the superiority of abstract elements in the analysis and benchmarking, transcendent. Scientific recognition and appreciation that is currently enjoying fuzzy techniques have led us to consider them useful in the training process.

Keywords: *training techniques, fuzzy, vague sets, evaluation, teaching process*

Introduction

1. Evaluation is part of the planning process of learning. Planning must be provided for teaching periods in which both teacher and student obtain and use information on changes proposed by the teaching purpose. Evaluation should be designed flexibly to meet both initial requirements and ideas that arise during the evaluation process. Planning should include strategies to ensure that students understand the lesson objectives and assessment criteria. Evaluation process has planned how students will receive their assessment results and how to participate at all stages of evaluation.

2. The assessment must follow the way students learn. Both student and teacher should think of learning when an assessment is scheduled and how it is interpreted outcome. Students should be aware of how they learned from all you have traveled.

3. Technical evaluation must be recognized as the principal means of interaction of the class.

Much of what the teacher and student interaction is reflected in the assessment. Questions, feedback from student demonstrates the level at which it arrived. Mode of expression and how it is often follows the observed and analyzed, being issued judgments on how to improve the course. Therefore, the assessment process adds value to the educational process and is a common practice in education.

4. Assessment should be seen as an important professional skill of the teacher.

Teachers make use of personal knowledge and experience of: planning evaluation, learning observation, analysis and interpretation of assessment results, communication of assessment results to students and to form habits of self.

Contents

The literature presents series of algorithms derived from fuzzy set theory. In this paper we have chosen and treated as a mathematical model developed by Professor Radu Ioan Academy of Economic Studies, Managerial and IT disciplines holder of the e-Organization, stating that this algorithm has been tested both theoretically in many graduation papers, theses and articles published in prestigious journals and practical research in national and international contracts.

In short, IT and management work coordinated by Professor Radu stated that based on models developed by R. and subsequently Yagger GD Zimmerman, occurred a series of algorithms developed and widely used especially by Japanese management, decision-making factors that associate membership criteria optimal variant in the interval [2.6].

The value associated with this factor must meet the following rule: if the criterion "j" is, in making the decision, how important criterion "k", then the coefficient associated with membership will have a value of closer to 6 than the corresponding coefficient criterion "k".

Starting from the matrix and the relative consequences of membership coefficient values, these algorithms require completion of a series of sequential steps and application of specific methods in the end that reflects the degree of optimism / pessimism of the decision maker.

The first step requires that the relative values of the absolute consequences. This involves generating a matrix, denoted by $CR=(cr_{ij})$ *cu* $i=1...,n$ *şi* $j=1...,m$**.** A certain element is obtained from the relationship:

$$\mathbf{Cr_{ij}} = \frac{\mathbf{Ca_{ij}}}{\sum_{i=1}^{n} \mathbf{Ca_{ij}}} \qquad (1)$$

where:

$\mathbf{Cr_{ij}}$	=	relative consequence of the alternative decision and the criterion j
$\mathbf{Ca_{ij}}$	=	absolute consequence of the alternative and criterion j
$\sum_{i=1}^{n} Ca_{ij}$	=	is the sum of the absolute consequences of alternatives in a given criterion

Formalizing the problem requires the expression of computational algorithms in a form that allows its subsequent formalization by computer-supported language. It also aims to ensure the character of generality, to increase the possibility that the software developed to solve any similar problem in terms of parameterization.

The second stage involves the calculation of matrix elements supporting Z (N, M). This requires the choice of the criterion of maximum or minimum, which is calculated based on cj, which is most favorable relative consequence of criterion j.

N	Variable memory, natural value, which means the number of decision alternatives
M	Variable memory, natural value, which indicates the number of criteria
CA(N,M)	N matrix of the absolute consequences of alternative criteria relative to M
K(M)	Vector of coefficients belonging to each criterion in the optimal variant
CR(N,M)	Relative incidence matrix
S(M)	Vector of absolute values of all the consequences of alternatives for each criterion, amounts necessary to calculate matrix elements of CR (N, M)

In vector CS (M) will hold the most favorable relative consequences for each of the M criteria. Calculation of matrix elements supporting Z = (zij), with i = $\overline{1,n}$ j = $\overline{1,m}$ and is using the expression:

$$Zij = \frac{|cr_{ij} - c_j^*|}{c_j^*} \quad (2)$$

As a consequence of the relatively favorable result means that the lowest relative criterion is the minimum that most if the criterion is maximum.

The third stage involves the matrix elements characteristic functions, this matrix CF = (fcij) and with i = $\overline{1,n}$ j = $\overline{1,m}$

A feature according to some criterion fcij associated with alternative i and j is expressed as:

$$fc_{ij} = e^{-k_j z_{ij}} \quad (3)$$

where:

kj = coefficient assigned to criterion j.

In the fourth stage is determined the optimal decision-making alternative. It will combine the functions of the characteristic matrix method of making a decision.

In this category of specific methods include:

a) Pessimistic technique (Abraham Wald), which is considered the best option to ensure obtaining the maximum benefit when presenting the objective conditions are unfavorable. Mathematically, the optimal choice is determined using the formula:

$$\textbf{V optimal} = \max_i \min_j FC_{ij} \quad (4)$$

b) Optimistic technique aimed at the best alternative where conditions appear most favorable targets, using the formula:

$$\mathbf{V\ optimal = max_i max_j FC_{ij}} \quad (5)$$

c) Technical proportionality (Bayes-Laplace) assumes that each state has the same objective conditions likely to occur and is optimal alternative to the arithmetic mean is most characteristic features. Mathematically, the formula for determining the optimal alternative is:

$$\mathbf{V_{optimala} = max \frac{1}{n} \cdot \sum_{j=1}^{n} FC_{ij}} \quad (6)$$

d) Technical optimality (Leonid Hurwicz) is a variant of the pessimistic and optimistic techniques, focusing on introducing a coefficient of optimism α (0 <α <1). Use the formula:

$$\mathbf{V\ optimal = max\ [\alpha \cdot Ai + (1-\alpha) \cdot ai]} \quad (7)$$

where:

ai = worst element of the alternative i

Ai = most favorable element of alternative i

Determine the optimal alternative can be using one of the above methods or by using several methods and classification decision alternatives according to each of them, then the results will be compared. The theory of vague sets given above, the basic concepts are: selection criteria and decision alternatives.

Students are alternatives and the selection criteria decision tasks can be accomplished. Each student is required to fulfill all tasks. How to perform the tasks of evaluation is essential.

Assuming that a student has achieved the following tasks: acquisition of concepts in a limited period of time and integrating them into new contexts, then, for evaluation, you can use traditional methods such as multiple choice test and essay. In Table 1 are the results obtained by students.

Table 1

Criteria / Alternative	Grid points obtained in the test	Time for delivery of three tasks	Number of concepts re-contextualize
Student 1	90 points	90 min	5 concepts
Student 2	80 points	110 min	7 concepts
Student 3	85 points	75 min	6 concepts

Note that the results are balanced, each of the students falling within the maximum time considered to be 2 hours and achieving a minimum performance set at 50 grid points and three test concepts / essay. To rank students according to performance, most often used in the formulation of conventions. Thus each task can be marked with a number of conventional points. For example 120 minutes can turn into a conventional point, 100 minutes in two points and so conventional The same concepts can proceed and if placed in new contexts: three integration concepts in the essay - one point, four concepts - 2 points etc. Signifies both the number of points obtained in task performance and share tasks in usability evaluation. But just here comes subjectivity. What are those reasons showing how many points to reward Evaluator with each activity. Such situations occur often in practice when a call for tenders and define the criteria or evaluate staff on a score sheet with set arbitrarily. Reasoning based on fuzzy set theory comes to overcome these suspicions. If the first step we identify the input variables at the second step discusses these variables. Because students will be defined not by their natural attributes, but through the selection criteria, these criteria we will comment further.

The first observation on the criteria identified in this case study called or generic criteria for quantification is that we need a unit and a scale of values. In most cases, the preferred standard measurement units are: meter length, gram for mass, volume liters or cubic decimetre etc. The fact that there are already instruments for measurement and verification for these units is an advantage. The drawback of these units is that in many situations due to the heterogeneity of the criteria, they differ and can hardly be interchangeable. In the previous example, for the grid points are used: minute for time unit and accepted for contextualization is the number of concepts.

Another observation would be generated when all the criteria could be measured quantitatively by the same measure, be it and bad. A constraint in sizing criterion that gove a favorable or unfavorable nature of its decision to the alternative. We believe that a test is positive if the alternative assessment unicriterial decision is made strictly according to its quantitative dimensions. In other words, the more quantitative benchmark size is larger, consider alternative more favorable decisions. Conversely, if a test is negative then the alternative decision criterion is smaller size as we consider alternative more favorable decisions. For example, the score awarded to an assessor criteria grid is considered favorable. Thus, higher the score is greater than consider alternative decision criterion is characterized by more favorable. Project duration of pregnancy is considered a criterion evaluator alternative unfavorable decision task because he wants to achieve in a short time. The period of performance is lower, we will consider alternative decision criterion is characterized by more favorable.

A third observation about the nature of the criteria points out that there are criteria that allowed quantification quantitative measurement tool using the same scale of values and, implicitly, the same unit. We believe that these criteria have the same characteristic favorable (or unfavorable) relative to alternative decisions. Yet its influence decision-making in selecting alternative may differ significantly indicating the level of importance given to the evaluator criterion. For example, the assessor may consider the test score scale should have a share of 30% in the final scoring.

Other observations about the nature of dependence criteria may be considered an alternative decision criterion, the compulsory or voluntary criteria, implications, or in other criteria, etc. In order not to complicate too many case studies we discuss limited to only three observations on the nature of the criteria:

the unit, the type of criterion (positive or negative) and level of importance. In Table 2 summarizes the debate over the nature of the criteria.

Table 2

Criterion	Weight	Unit Type	Level of Significance Criterion	Other Elements
1. Test scores obtained at grid points	points	Favorable	30%	
2. Project duration of the tasks	minutes	Unfavorable	10%	
3. Re-contextualize concepts	piece	Favorable	60%	

Reviewing the three observations (the number is not exhaustive) criteria on the nature causes us to ways to clarify issues raised.

The first observation draws attention to the metrics of quantitative scale allowed for the identification of decision alternatives. Replacing them with conventional units (points) is not recommended because it raises questions about the rationale behind the establishment of weights and ranges used.

Economic and technical literature abounds with recommendations on the characterization of objects of knowledge by sizing their qualitative and not quantitative. For example, profit can be expressed quantitatively as evidenced by the size of conventional units and monetary units. If we inform a person that a company is profit 30 000 monetary units, that person can not decide whether or not to buy shares in this company because the size called quantitative characterizing income subject knowledge is not sufficient to reach a point The rationally.

You will need to communicate to that person and other quantitative dimensions of the objects of knowledge which led to obtaining the results such as turnover or revenue and expenditure, although the quantitative dimension of profit (the result) is obtained from the difference between the two variables. But if we consider a profit function (dependent variable) by two independent variables (turnover and costs) when we seek the outcome of which algorithm to store the knowledge of factors and elements. Putting the two independent variables as a ratio (fraction) give us another image of the object attribute profit by his knowledge, known rate of return.

We believe that the profit rate is an indicator because if that organization issued shares on the stock market, investors may perceive the risk.

Some authors refer to the relationship between two quantitative indicators (one on one effort effects, or vice versa) qualitative indicators. Qualitative indicators are expressed in percentage units and not units. This feature allows us to provide quality indicators with their tasks without the need of conversion.

From Table 1 which includes the selection criteria of trainees will retain only the actual values recorded (Table 3). In the literature these values is called the structure matrix specifying absolute consequences that the matrix can contain only numeric values is necessary to express a numerical scale of the consequences of absolute values.

Table 3

Criteria / Alternative	C1	C2	C3
A1	90	90	5
A2	80	110	7
A3	85	75	6
A4	100	120	4
Suma	*355*	*395*	*22*

Next we solve this problem by following the sequence of stages of decision theory specifies vague sets, accompanied by commentaries.

Step 1. Calculation of matrix elements relative consequences of CR (4.3)

Step aims to transform the absolute values to quantify the relative values of the criteria considered in the selection of students (Table 4.). On absolute matrix elements consequences can use the formula:

$$Cr_{ij} = \frac{Ca_{ij}}{\sum_{i=1}^{n} Ca_{ij}}$$

Table 4

Criteria / Alternative	C1	C2	C3
A1	0,25	0,23	0,23
A2	0,23	0,28	*0,32*
A3	0,24	*0,19*	0,27
A4	*0,28*	0,30	0,18
c_j^*	**0,28**	**0,19**	**0,32**

Step 2. Calculation of matrix elements supporting Z (4,3)

Formula as supporting matrix elements are calculated as follows:

$$Zij = \frac{\left|cr_{ij} - c_j^*\right|}{c_j^*}$$

where Cj * is the consequence of favorable relative criterion j (if the test is considered positive when at least the smallest consequence that most if the criterion is maximum. The table below supporting matrix elements are calculated.

Table 5

Criteria / Alternative	C1	C2	C3
A1	0,09	0,20	0,29
A2	0,20	0,47	0
A3	0,14	0	0,15
A4	0	0,20	0,18
k	0,30	0,10	0,60

Step 3. Calculation of matrix elements characteristic functions FC (4,3)

At this stage the technique applies fuzzy supporting elements according to the coefficient matrix of importance given by the operation:

$$fc_{ij} = e^{-k_j z_{ij}}$$

Table no. 6. calculate matrix elements characteristic functions.

Table 6

Criteria / Alternative	C1	C2	C3
A1	0,97	0,98	0,84
A2	0,99	0,95	1,00
A3	0,96	1,00	0,91
A4	1,00	0,94	0,90

Step 4. Establishing the optimal alternative

The establishment of optimal alternatives apply the matrix elements characteristic functions, balance technique (Bayes-Laplace). Balance technique is used in the formula:

$$Vopt. = max. \frac{1}{n} * \sum_{j=1}^{n} FC_{ij}$$

The values obtained are presented in Table 7.

Table 7

Criteria / Alternative	C1	C2	C3	$\frac{1}{n} * \sum_{j=1}^{n} FC_{ij}$
A1	0,97	0,98	0,84	0,93
A2	0,99	0,95	1,00	0,98
A3	0,96	1,00	0,91	0,96
A4	1,00	0,94	0,90	0,95

A graphical representation is:

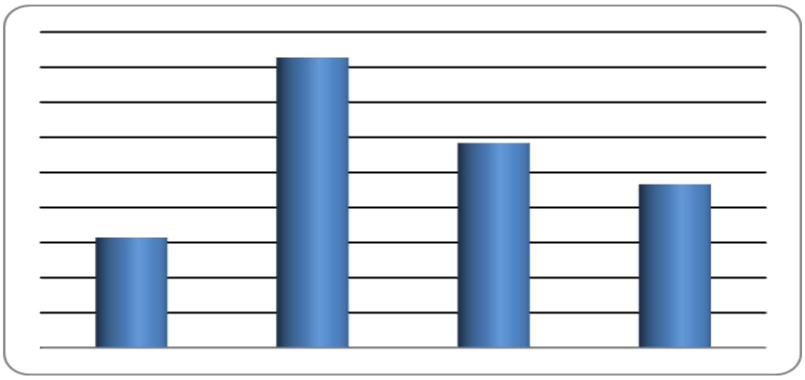

Fig. 1. Alternative set of decision-making hierarchy technique Bayes-Laplace

Conclusions

Objectivity of scoring students has been and will remain a subject of dispute in education. Identifying methods that allow the learner acceptance of a classification is the desire of every teacher. Whichever method of valuation used, or the science of scoring docimologic, recommended that the final distribution of grades per student to take the form of the Gauss curve (graphical representation of the probability distribution of mean values according to a standard).

In other words, bring to your attention that any vote, which ultimately does not lead to a distribution that is as close as the Gauss, is prone to interpretation. We present two conflicting situations often encountered in practice:

1. Grouping results on a subinterval (eg between 6 and 8) and not scoring range (in education in Romania notes are from 1 to 10) draws a number of criticisms of the scoring because, although the group of students is homogenous, the items used in scoring did not allow for sufficient refining peculiarities. It is recommended that appropriate weighting results to reach the recommended range.

2. Distribution, almost uniform, the performance of students in the recommended range can be interpreted as highlighting the features and elements subject to assessment could be considered the strong point of this notation. However, and this rating may be subject to criticism considering that the heterogeneity of students and the results could be due to external factors, uncontrollable (eg, charging different concepts depending on bibliographic resources accessed by students, some of which are accepted scientific community).

If the problem is to obtain outcomes identified distributed as "Gauss," then one solution may be assessing the comparability. This involves correcting and grading works according to the maximum level of performance achieved in a paper. For example, you read all the entries and select the best and worst work. They will get the maximum score that minimum. Scoring items are set and distributed so as to obtain notes Gauss.

Although at first glance, the proposed solution (the comparability assessment) appears to exceed the above statements (or subinterval grouping nearly uniform distribution throughout the range), it nevertheless raises questions about the standard of quality that may relate to the whole group. It may happen that, although none of the papers assessed did not meet the minimum requirements to obtain the maximum score, or the opposite, even if all works fully meet the requirements, some of them to obtain lower scores. This raises the problem of defining requirements to enable earlier assessment evaluator to shape a picture of the student performing the task.

The test grid is perceived by most students, as an objective means of scoring. Criticism of this form of notation, in particular due to the wording of questions and their difficulty level. It is considered that the appearance of a large number of students of unsatisfactory results were due by the teacher evaluator unidentified the average level of knowledge to formulate appropriate requirements. Identifying the average level can lead to surprising results obtained favorable perception of the teacher which would distort the level of knowledge by learners. This is because one of my roles is assessing distinctiveness, identifying those features that allow the teacher to decide future action on ways suited to each student.

This raises the question whether science is able to respond to such challenges. Addressing the vague set theory I tried to formulate a possible answer rationally acceptable.

Bibliography

1. **GOTTWALD** Siegfried,. *A Treatise on Many-Valued Logics*. Baldock, Hertfordshire, England: Research Studies Press Ltd., 2001
2. **MÖLLER**, B, and **BEER**, M - *Fuzzy Randomness - Uncertainty in Civil Engineering and Computational Mechanics*, Springer, Berlin Heidelberg New York. 2004
3. **RADU, Cătălin** – *Evaluarea on-line - valori, metodologii, referinţe,* Universitatea Naţională de Apărare „Carol I”, Conferinţa Naţională de Învăţământ Virtual, ediţia a VI-a, 2008
4. **RADU Ioan, BURLACU Sorin** ş.a.- *Informatică şi Management* **-** *o cale spre performanţă,* Editura Universitară, Bucureşti, 2007
5. **RADU, Ioan** (coordonator) – *Fundamentarea deciziilor complexe prin tehnici de simulare* – Editura Economică, 2002
6. **RADU, Cătălin** – *Evaluarea on-line - valori, metodologii, referinţe,* Universitatea Naţională de Apărare „Carol I”, Conferinţa Naţională de Învăţământ Virtual, ediţia a VI-a, 2008
7. **SHARON, Y.** (EDT) Tettegah - *Technology and Education* – Elsevier, Hardcover, 2006