Computers & Industrial Engineering 57 (2009) 732–741

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

An adaptive neuro-fuzzy model for prediction of student's academic performance ☆

Osman Taylan [a,*], Bahattin Karagözoğlu [b]

[a] Department of Industrial Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia
[b] Department of Electrical and Computer Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Article history:
Received 6 November 2007
Received in revised form 18 November 2008
Accepted 30 January 2009
Available online 5 February 2009

Keywords:
Neuro-fuzzy system
Student academic performance
Learning fuzzy models

ABSTRACT

This paper introduces a systematic approach for the design of a fuzzy inference system based on a class of neural networks to assess the students' academic performance. Fuzzy systems have reached a recognized success in several applications to solve diverse class of problems. Currently, there is an increasing trend to expand them with learning and adaptation capabilities through combinations with other techniques. Fuzzy systems-neural networks and fuzzy systems-genetic algorithms are the most successful applications of soft computing techniques with hybrid characteristics and learning capabilities. The developed method uses a fuzzy system augmented by neural networks to enhance some of its characteristics like flexibility, speed, and adaptability, which is called the adaptive neuro-fuzzy inference system (ANFIS). New trends in soft computing techniques, their applications, model development of fuzzy systems, integration, hybridization and adaptation are also introduced. The parameters set to facilitate the hybrid learning rules for the constitution of the Sugeno-type ANFIS architecture is then elaborated. The method can produce crisp numerical outcomes to predict the student's academic performance (SAP). It also provides an alternative solution to deal with imprecise data. The results of the ANFIS model are as robust as those of the statistical methods, yet they encourage a more natural way to interpret the student's outcomes.

1. Introduction

Fuzzy systems are fuzzy model structures in the form of fuzzy rule bases (FRBs) that are the most important area in the application of the fuzzy set theory. Designing a fuzzy rule based system involves derivation of the desired 'If-Then' fuzzy rules, partitioning of universes, and addressing of the membership functions (Nie, 1995).

A few systematic design procedures are available for the development of learning models even though fuzzy systems have reached a recognized success in several significant application areas such as control problems (Magdalena & Monasterio, 2000), modeling (Papadakis & Theocharis, 2002), performance classification (Ishibuchi, Nakashima, & Murata, 1999; Taylan, 2006), and decision systems (Lin & Lee, 1991). The trial and error approach is a natural choice to design fuzzy systems since their originations. In trail and error approach, fuzzy rules can be easily and directly formulated by experts in the form of linguistic rules, yet the fuzzy system performance does not suffer from critical degradation due to parameters defined in a non-optimal way. Nevertheless, this approach is neither suitable nor feasible when there is no linguistic knowledge available (Nie, 1995; Pedrycz, 1993). Many researchers have put efforts to develop alternative design methods. For instance, clustering techniques are considered as plausible approaches for deriving fuzzy rules (Figueiredo & Gomide, 1999) albeit methods based on clustering usually do not provide fuzzy systems described in the linguistic form. In addition, many design methods based on qualitative approaches and optimization theory have also been suggested. However, local solutions and slow convergence are very common in algorithms based on quantitative methods, which often impose practical constraints in using the corresponding fuzzy methods (Jang & Sun, 1995). Manual design of fuzzy systems may require a long period of trial and error and much input from experts due to the difficulty in defining the choice of membership functions or reasoning structures.

Fuzzy systems lack learning ability. What may be adequate for one case with a given set of conditions may not be appropriate for another case that looks similar but has a different set of conditions. To improve the performance of a fuzzy system, the parameters in terms of the structure, complexity, type of networks, etc., have to be tuned and determined. In recent years, it has become clear that neural and fuzzy hybrid systems have advantageous such as the applicability of existing algorithms for artificial neural networks

☆ This manuscript was proposed by Area Editor D.L. Kimbler.

* Corresponding author. Tel.: +966 500031056; fax: +966 26952486.
E-mail addresses: otaylan@kau.edu.sa (O. Taylan), bkaragoz@kau.edu.sa (B. Karagözoğlu).

0360-8352/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cie.2009.01.019

(ANNs), and direct adaptation of knowledge expressed as a set of fuzzy linguistic rules (Chowdhury & Li, 1998).

In this paper, we propose a class of neuro-fuzzy networks with the ultimate aim that is to design a fuzzy inference system (FIS) via learning. The underlying network structure and learning algorithm imply a systematic approach for the FIS design. Both neural networks and fuzzy systems are dynamic, parallel processing systems that estimate input–output functions. The main features of the proposed approach can be briefed as:

(1) it processes data according to fuzzy reasoning mechanisms,
(2) it recovers the encoded knowledge in the form of fuzzy rules and put them in the usual,
(3) the desired number of rules to represent the FIS is the structural decision, the input space partition and parameters are being automatically determined by the learning algorithm,
(4) it learns rules covering the whole input space and it has universal approximation character,
(5) it processes the data and automatically determines the membership functions ranges.

Details of the approach are given in Section 4. The network learns in two main phases. In the first phase, the network self organizes to determine the rules and the respective antecedent membership functions from the data set. In the second phase, it learns consequent parameters using a supervised scheme (Lin & Lu, 1996).

The linguistic variables and their labels are the backbone of fuzzy rule base systems and the fuzzy models. Schmucker (1989) has defined a linguistic variable as a variable whose values are words or sentences in a natural or synthetic language. Such statements as '*He has learnt very well and did excellent in the final exam*' are really vague linguistic terms in real life since someone may ask; '*how very well?.*' The ANFIS under consideration is a multi-input single-output (MISO) system which has six inputs and one output linguistic attribute. The input attributes of the ANFIS system are the assessment tools that are *Quiz (Q), Major (M), Midterm (MD), Final (F), Performance appraisals (P),* and *Survey (S)'* and the output is the '*Student's Academic Performance (SAP)'*. These imprecise attributes are called fuzzy linguistic variables and used commonly in an educational system. They are introduced and expressed by premise parameters (fuzzy linguistic labels) in daily life such as '*unsatisfactory* (A_1)*, average* (A_2)*, good* (A_3)*, very well* (A_4)*,* and *excellent* (A_5)*.* Arithmetical and statistical methods are unable to offer an effective inference procedure to perform the evaluation of the academic performances of students in a more natural way, using linguistic variables. They usually use numerical values to express the students' achievements. Law (1995) expresses the natural linguistic terms (*e.g. unsatisfactory, average, good, very good, excellent etc.*) to represent the SAP by ignoring mostly their inherent nature of vagueness. However, the academic performance evaluation involves the measurement of ability, competence and skills that are uncertain and imprecise concepts captured in fuzzy systems. The use of a natural language such as "*good, very good, and excellent*" would yield more meaningful explanations and could allow flexibility in reasoning and judgment of the student's performance.

The current study focuses on development of a data-driven adaptive neuro-fuzzy system (ANFIS) using a real dataset obtained from the last 4-year period of students' achievements in the engineering economy course. This method might help students, their parents, decision makers, and evaluators in obtaining more reliable and understandable results for a student's achievement, or for a group of students and their comparative evaluations. It is important to point out that the aim of proposed method is not to replace the traditional method of evaluation; instead, it is to strengthen the present system by providing additional information for decision making.

The paper is organized in five sections. After the introduction in Section 1, Section 2 describes neuro-fuzzy systems integration, hybridization and adaptation. Section 2 continues with explanations of genetic fuzzy systems and adaptive neuro-fuzzy systems (ANFIS). Assessment of the student's academic performance (SAP) is addressed in Section 3, which also introduces the existing methods of the performance evaluation. Section 4 discusses the development of ANFIS, the fuzzy reasoning procedure, fuzzy rule sets and membership functions. It continues with discussions on the architecture of hybrid learning and fuzzy model validation, the error of observations for checking and training data sets. Section 5 presents the conclusions of the research, and the outcomes of SAP for Statistical and ANFIS methods for various observations. The paper ends with a list of references.

2. Neuro-fuzzy integration, hybridization and adaptation

Fuzzy systems and Artificial neural networks)ANNs) are soft computing approaches for modeling human (domain expert) behaviors. The goal is to mimic the actions of a domain expert who can solve complex problems. Hence, a learning process can be a knowledge acquisition system in the absence of a domain expert or insufficient time or data. On the other hand, one can build a fuzzy system if he/she has knowledge expressible in linguistic forms. The relationship between neural networks and linguistic knowledge based system is bidirectional and broadly discussed in (Ishibuchi, Nii, & Turksen, 1998). Neural network-based systems can be trained by numerical data and fuzzy rules can be extracted from neural networks (Ishibuchi et al., 1998). Similarly, fuzzy rule based classification systems can be designed by the linguistic knowledge. The fuzzy logic and neural systems have very contrasting applications requirements too. For instance, fuzzy systems are appropriate if sufficient expert knowledge about the process is available, while neural systems are useful if sufficient process data are available or measurable. Both approaches build nonlinear systems based on bounded continuous variables, the difference being that neural systems are treated in a numeric quantities manner, whereas fuzzy systems are treated in a symbolic qualitative manner (Mitra & Hayashi, 2000). Therefore, the integration of neural and fuzzy systems leads to a symbolic relationship in which fuzzy systems provide a powerful framework for expert knowledge representation, while neural networks provide learning capabilities. The aim of integration is to build more intelligent decision making systems.

There is an increasing interest to augment fuzzy systems with learning and adaption capabilities. This enhancement effort is succeeded by hybridizing the approximate reasoning method of fuzzy systems with learning capabilities of neural networks and genetic algorithms. The merits of neural networks and fuzzy systems can be integrated in a neuro-fuzzy approach (Pal & Mitra, 1999). Eventually, the combination of neural networks and fuzzy systems (neuro-fuzzy systems) has been recognized as a powerful alternative approach to develop adaptive soft computing systems. Consequently, two main approaches of hybridization attempts flourish in the framework of soft computing as the neural networks providing fuzzy systems with learning capabilities (NFSs) and genetic fuzzy systems (GFSs).

Neuro-fuzzy systems are one of the most successful and visible directions of that effort. Neuro-fuzzy hybridization is done in two ways (Mitra & Hayashi, 2000): a neural network equipped with the capability of handling fuzzy information (termed fuzzy neural network) and a fuzzy system augmented by neural networks to enhance some of its characteristics like flexibility, speed, and adaptability (termed neuro-fuzzy system(NFS) or ANFIS). An adapted neuro-fuzzy system (NFS) is designed to realize the process of fuzzy reasoning, where the connection weights of network correspond to parameters of fuzzy reasoning (Berenji & Khedkar, 1992; Mitra & Hayashi, 2000). These methodologies are thoroughly discussed in the literature (Mitra & Hayashi, 2000).

Fuzzy systems can be broadly categorized into two families. The first category includes linguistic models based on collection of If-Then rules, whose antecedents and consequences utilize fuzzy values. It uses fuzzy reasoning and the ANFIS (Mamdani, 1997) model falls into this group. However, the added complexity in structure and computation of the ANFIS with max–min composition does not necessarily imply better learning capability or approximation power. The second category bases on the Sugeno-type system (Jang, Sun, & Mizutani, 1997) that uses a rule structure with fuzzy antecedent and functional consequent parts (Mitra & Hayashi, 2000). In this study, we concentrate on the ANFIS architecture for the first order Sugeno fuzzy model because of its transparency and efficiency. We also describe how to decompose the parameter set to facilitate the hybrid learning rule for adaptive neural fuzzy inference system (ANFIS) architecture representing the Sugeno fuzzy models.

2.1. Genetic fuzzy systems (GFSs)

A second and distinct approach to hybridization is the genetic fuzzy systems (GFSs) (Cordon, Gomide, Herrera, Hofmann, & Magdalena, 2004). A GFS is essentially a fuzzy system augmented by a learning process based on genetic algorithms (GAs). GAs are search algorithms, based on natural genetics, that provide robust search capabilities in complex spaces. Genetic learning processes cover different levels of complexity according to the structural changes produced by the algorithm, from the simplest case of parameter optimization to the highest level of complexity of learning the rule set of a rule based system. The parameter optimization has been the approach used to adapt a wide range of dissimilar fuzzy systems, as in genetic fuzzy clustering or genetic fuzzy systems (Cordon et al., 2004). However, genetic fuzzy systems are not subject of this work. They might be considered for the estimation of student academic performance in future.

2.2. The adaptive neuro-fuzzy inference system (ANFIS)

An adaptive network is a neuro-fuzzy system (ANFIS) with an overall input–output behavior that is determined by collection of modifiable parameters. The structure of an adaptive network is composed of nodes connected by directed links, where each node performs a function on its incoming signals to generate a single node output and each link specifies the direction of signal flow from one node to another (Jang et al., 1997). Usually, each node function is a parameterized function with modifiable parameters; by changing these parameters, the node functions as well as the overall behavior of the adaptive network, are changed. Each node in an adaptive network performs a static mapping from its input(s) to output. To facilitate the development of learning algorithms, it is assumed that all node functions are differentiable except at a finite number of points. The parameters of an adaptive network are distributed into its nodes with each node having a local set of parameters. The union of these local parameter sets constitutes the network's overall parameter set. The adaptive networks proposed in this study are functionally equivalent to fuzzy systems.

The proposed network structure encodes If-Then fuzzy rules in which the consequent outcome is a function of the input variables. The network structure estimates a function without any mathematical model to learn from experience with data samples. Fuzzy sets are considered to be advantageous in the logical field, and in handling higher order processing easily. The higher flexibility is also a characteristic feature of neural nets produced by learning that makes them more suitable to a data-driven processing. It has been proved that: any rule based fuzzy system may be approximated by a neural net or any neural net may be approximated by a rule based fuzzy system (Mitra & Hayashi, 2000). Jang and Sun (1993) have shown that fuzzy systems are functionally equivalent to a class of radial basis function networks, based on similarities between the local receptive fields of the network and the membership functions of the fuzzy system.

The basic drawback of the current ANFIS is the addressing of fuzzy linguistic parameters and their term sets that must be identified by the designer. In general, the designer chooses the types of membership functions. In many neuro fuzzy design methods, the fuzzy sets involved are defined in multidimensional spaces. The number of membership functions for each input space might be a problem for an ANFIS design besides the parametric learning and structure learning problems dealing with the partition on the input–output universes (Figueiredo & Gomide, 1999). In addition, the network structures are prefixed and may be appropriate only for a limited set of problems, and eventually the learning algorithm may be trapped in local optima. There is a main difficulty in obtaining effective and quality training data for engineering systems due to the fact that the generated training data set may not be predictive, thus it cannot truly represent the varying environment.

3. Review of the existing methods of performance evaluation

Education is a process which accepts students from a low level of standing, pushes them through several stages of development and produces individuals qualified with certain abilities, skills and attributes who are fit for a job or a higher level in education. Courses are indispensible ingredients of the process. Steps in the process are specified as the curriculum. Traditionally, the characteristics of the overall system have been defined by models that have been specified by the number of courses students have completed and grant point averages (GPA) they accumulated in them. This approach however, doesn't guarantee that students have achieved the qualifications required for the after era. Educationalists have stipulated a new approach which relies on what the student learns rather than what he/she has been taught. An act of assigning a qualitative or quantitative merit or worth to student achievements is defined as the academic assessment (Karagözoğlu & Türkmen, 2007). Assessment of the student's academic performance (SAP) is one of the most important practices used for three main reasons: to decide on pass and failure in courses, to obtain an indication of the student's level of learning, and to provide information on the effectiveness of teaching.

Graduates of an educational system will be implementing the knowledge and skills they have gained during their studies in work places. In the engineering field, the half-life is around 5 years; i.e. in 5 years time only half of what the student learns at the university remains useful. Expectations of stake holders from the graduates in this duration are defined as program educational objectives (PEO). Abilities, skills and attributes that must be gained by students at the graduation to fulfill the PEOs are stated as program outcomes (POs). Foundations of POs in courses are specified as course learning objectives (Karagözoğlu, 2007). In this respect, SAP is the most likely measure to verify achievement of course learning objectives, to expose the effectiveness of learning environment and to monitor the standards of the educational organization.

Traditionally, a student's academic performance (SAP) has been evaluated based on the marks collected by the student. This approach has severe limitations in satisfying the SAP measurements for an outcome-based education. Fuzzy rule based models rely on linguistic rules that are easy to define and use (Taylan & Karagözoğlu, 2007). However, they lack learning capabilities (Chowdhury & Li, 1998). Combinations of fuzzy rules with artificial neural networks and genetic algorithms shall provide the flexibility of the develop assessment tool to adapt into novel situations. The tool must be robust, yet easy to use.

3.1. Existing methods of performance evaluation

In traditional (statistical) methods, the student's academic performance (SAP) is evaluated based on the marks collected by a student. It can be classified into numerous categories such as single numerical scores usually referring to 100 percent, single letter grades (e.g. A, B, C, D, or F), nominal scores (e.g. 1, 2, 3…10), linguistic terms such as "*Fail*", or "*Pass*" or single grade-points from 0.00 to 4.00. As a part of this study, a weighted sum of assessment tools is used to calculate the numerical score of each student as follows: *Quiz (Q)* is 10%, *Major (M)* is 15%, *Midterm (MD)* is 20%, *Final (F)* is 40%, *Performance Appraisals (P)* is 10%, and *Survey (S) is* 5%. The total out of 100 indicates the student's academic performance (SAP).

Assessment techniques have been discussed widely in the literature. (Ratcliff, Arkin, & Dove, 1997) worked on qualitative and quantitative assessment techniques. Similarly, (Lopes, Lanzer, & Barcia, 1997) studied the qualitative techniques based on several indicators for teaching, research, and quality of a department. They also declared that the indicators of "quantities" have fuzzy meanings. (Deniz & Ersan, 2002) presented several ways in which student's performance data can be analyzed and presented for an academic decision making and academic decision support system. (Ma & Zhou, 2000) presented an integrated fuzzy set approach to assess the outcomes of a student learning. They exploited fuzzy set principles to represent the imprecise concepts for subjective judgment and applied a fuzzy set method to determine the assessment criteria and their corresponding weights. Their study aimed at encouraging students to participate in the whole learning process as well as providing an open and fair environment for assessment.

Ebel and Prade (1991) evaluated the academic performance in two distinct forms; formative assessment and summative assessment. It is important to give feedback to students, teachers and educational planners during the course of instruction. Hence, the formative assessment is conducted to monitor the progress of instruction through series of tests, quizzes, and observations. The summative assessment is carried out at the end of each instructional segment through tests and final examinations to provide information on how much the students have achieved their objectives. However, the assessment is a part of learning process and it is not something to be done after the instruction is completed. Also, written examinations at the end of the semester may not provide a complete picture of what the students have learnt. Preferably, a combination of both assessment methods might be used to provide a full coverage of important learning outcomes. Eventually, in an educational system, assessment tools may consist of series of tests, quizzes, portfolios, formal written examinations, individual assignments and course works, group works, observations, projects, publishable materials and oral presentations.

The commonly used technique for assessment of the academic performance relies on statistical methods that award numerical values or linguistic labels to a given piece of student's work. These values and labels have been frequently used to represent the student's achievement without exploiting any other alternative to double check the student's performance. The grade awarded by an evaluator could be an approximation only, because he/she typically assigns a numerical score to the student's work and usually does not consider a comparative evaluation and linguistic reasoning. This score might vary as done by different evaluators because of their sensitivities, experiences and the standards.

4. Development of adaptive neuro-fuzzy inference system (ANFIS)

Developing an ANFIS model to evaluate the student academic performance (SAP), it is necessary to take into consideration the scarcity of data and the style of input space partitions. For example, for a single–input problem, usually 10 data points are necessary to come up with a good model (Jang et al., 1997). In this study, 228 data instances are used for the development of the ANFIS model. The common approach is to divide the data set into training and checking data sets. The training data set serves in model building while the other one is used for the validation of the developed model. The ANFIS model is able to produce crisp numerical values and includes the following parameters:

(1) defining input and output variables by linguistic statements,
(2) deciding on the fuzzy partition of the input and output spaces,
(3) choosing the membership functions (MFs) for the input and output linguistic variables,
(4) deciding on the types fuzzy control rules,
(5) designing the inference mechanism and
(6) choosing a de-fuzzification procedure.

Fuzzy clustering approach is used to generate objective number of rules which are based on the clustering of input and output data sets, the level of fuzziness of clusters, and the membership functions. In this approach, the number of rules is usually equal to the number of output clusters regardless of the number of input variables. The subtractive clustering is an off-line clustering technique used in conjunction with radial basis function networks (RBFN) and fuzzy modeling. It assumes each data point as a potential cluster center and calculates a measure of likelihood defining the cluster center based on the density of surrounding data points. To generate an ANFIS structure, a cluster radius must be specified to indicate the range of influence of the cluster. Fig. 1 illustrates a special case where the data space is considered as a unit hypercube. Specifying a small cluster radius yields many small clusters in the data, resulting in a lot of rules. Let us consider a collection of n data points $\{x_1,..,x_n\}$ in an M-dimensional space. Since each data point is a candidate for cluster centers, the density measure at data point x_i is defined by Eq. (1).

$$D_i = \sum_{j=1}^{n} \exp\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right) \tag{1}$$

where r_a is a positive constant. The data points near the first cluster center x_{c1} will have significantly reduced density measures. After the density measure for each c data point is revised, the next cluster center x_{c2} is selected and all of the density measures for data points are revised again. This process is repeated until a sufficient number of cluster centers are generated. These cluster centers will be used as the centers for the fuzzy rules' premise in a zero-order Sugeno fuzzy model (Cordon et al., 2004; Jang et al., 1997; Taylan, 2006). For instance, assuming that the center for the ith cluster is c_i in an M-dimensional space, the c_i can be decomposed into two component vectors p_i and q_i, where p_i is the input part that contains the first N element of c_i; and q_i is the output part that contains the last M-N elements of c_i. Then, for a given input vector x, the degree to which fuzzy rule i, if fulfilled, is defined by Eq. (2). In this study, five cluster centers have been determined for the given 228 data set. The number of fuzzy rule set would be equal to the number of cluster centers, each representing the characteristic of the cluster as given in Fig. 1.

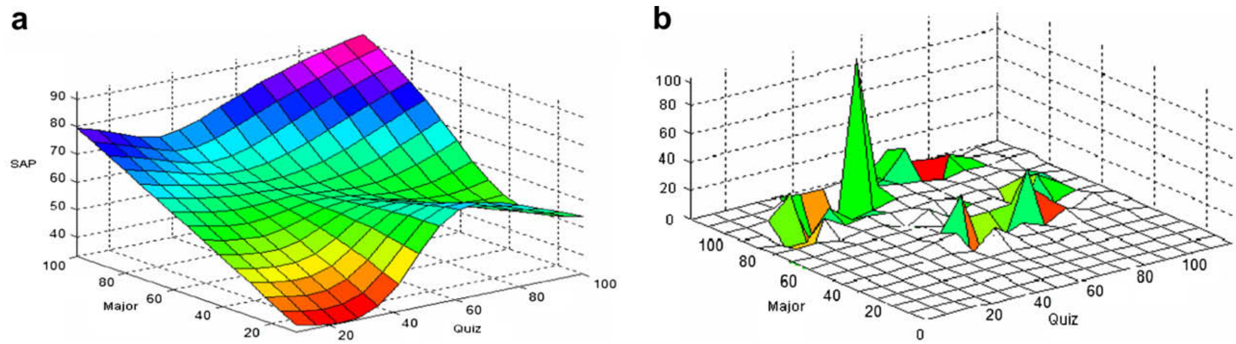

Fig. 1. Control action surface after training (a), and comparison of output (SAP) surfaces constructed by the five MFs, after (b) normalization for inputs (Major and Quiz).

$$\mu_i = \exp\left(-\frac{\|x - p_i\|^2}{(r_a/2)^2}\right) \tag{2}$$

4.1. Fuzzy reasoning, fuzzy rules and membership functions

Fuzzy rules and fuzzy reasoning are the backbone of fuzzy inference systems, which are the most important modeling tools based on fuzzy sets (Jang et al., 1997). Fuzzy reasoning is an inference procedure that derives conclusions from the set of fuzzy If-Then rules and known facts. Fig. 2 shows the reasoning procedure for a first order Sugeno fuzzy model. Since each rule has a crisp output, the overall output is obtained via a weighted average, thus avoiding the time consuming process.

The input parameters of the ANFIS under consideration are *Quiz (Q), Major (M), Midterm (MD), Final (F), Performance Appraisals (P)*, and *Survey (S)*' and the output is the '*Student's Academic Performance (SAP)*'. These imprecise attributes are called fuzzy linguistic variables are used in any educational system. These linguistic variables are imprecise, vague and incomplete fuzzy terms. They are introduced and expressed by fuzzy linguistic values such as '*unsatisfactory* (A_1), *average* (A_2), *good* (A_3), *very well* (A_4), *excellent* (A_5)', as is given in Fig. 3.

$\mu_A(x)$, where x is the mark to represent an achievement of a student in set A. Fuzzy rules are mathematical relationships mapping the inputs to output relations and they are constituted from fuzzy linguistic variables and their term sets. Fuzzy If-Then rules are known as fuzzy implications or fuzzy conditional statements that are widespread in our daily linguistic expressions. Fuzzy rules are the backbone of an ANFIS model. For example; '*IF the Quiz (Q) of a student is Good and Major (M) exams are Good and Midterm (MD) exam is Very-good and Final (F) exam is Good THEN The SAP will be Good*' is a complete rule defining the relations of input and output linguistic variables.

$$\begin{aligned}
&\text{Rule } 1: \text{ If } Q \text{ is } A_1 \text{ and } M \text{ is } A_2 \text{ and } \ldots \text{ and } F \text{ is } A_2 \\
&\qquad \text{Then } f_1 = p_1 Q + q_1 M + \cdots + m_1 F + r_1 \\
&\text{Rule } 2: \text{ If } Q \text{ is } A_2 \text{ and } M \text{ is } A_4 \text{ and } \ldots \text{ and } F \text{ is } A_5 \\
&\qquad \text{Then } f_2 = p_2 Q + q_2 M + \cdots + m_2 F + r_2 \\
&\quad \cdots\cdots\cdots \qquad \cdots\cdots\cdots \\
&\text{Rule } n: \text{ If } Q \text{ is } A_n \text{ and } M \text{ is } A_n \text{ and } \ldots \text{ and } F \text{ is } A_n \\
&\qquad \text{Then } f_n = p_n Q + q_n M + \cdots + m_n F + r_n
\end{aligned} \tag{3}$$

As was pointed by Zadeh (1972), conventional techniques for system analysis are not suited for dealing with a humanistic system, whose behavior is strongly influenced by human judgment, perception, and emotions. This belief yields the concept of linguistic variables as an alternative approach to model human thinking.

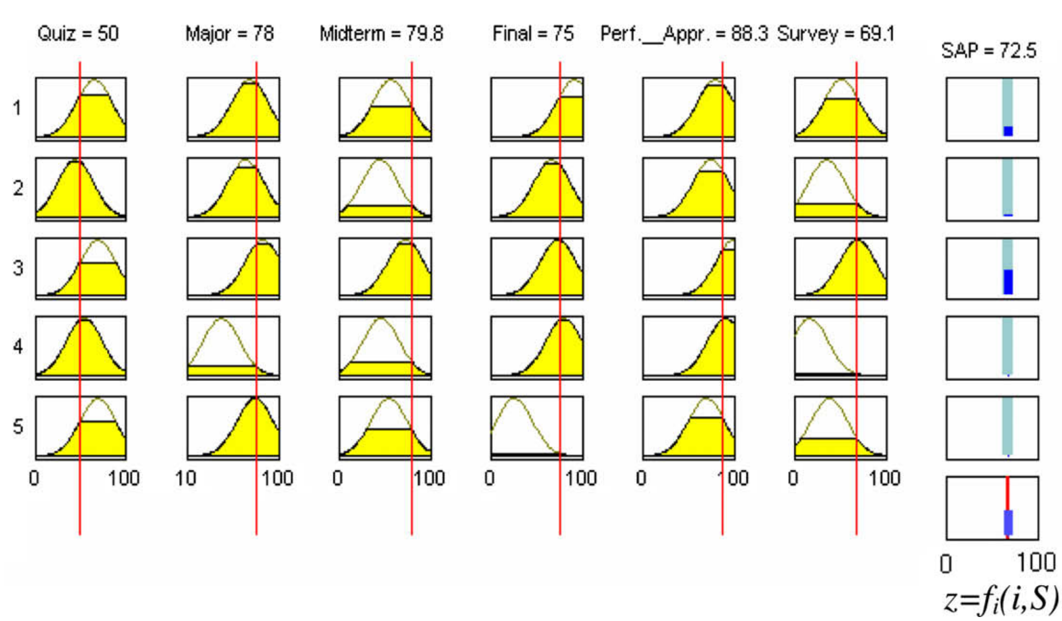

Fig. 2. Fuzzy reasoning procedure for Sugeno model of SAP assessment.

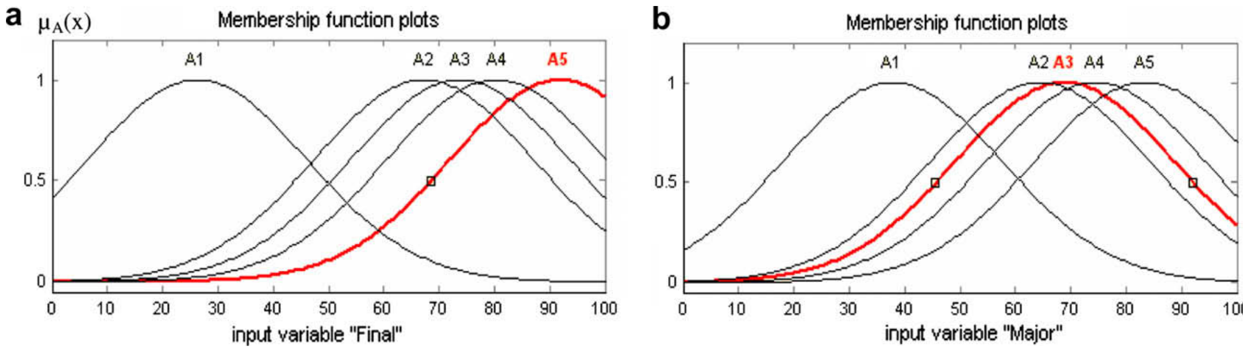

Fig. 3. Fine tuned membership functions of 'Final exam (a), Major exam (b)' as input variables.

The rule set in Eq. (3) illustrates the reasoning mechanism and the corresponding equivalent ANFIS architecture where the nodes of the same layer have similar functions.

Mendel (2001) points out that the fuzzy If-Then rules are production rules including antecedent parts, and consequence part. A typical fuzzy rule in a Sugeno fuzzy model has the form in Eq. (3), where $A_1,\ldots,A_5$ are fuzzy linguistic values, defining the linguistic variables. *'Q is A_1' or 'M is A_2'* are antecedents. $z = f_i(x,y)$ is a crisp function in the consequent and the results of this function are also crisp values. Usually, $z = f_i(x,y)$ is a polynomial in the input–output variables x and y, but it can be any function as long as approximately describing the output of the model within the fuzzy region specified by the antecedent of the rule. When $z = f_i(x,y)$ is a first order polynomial, the resulting ANFIS is called a first order Sugeno fuzzy model (Berenji & Khedkar, 1992). Moreover, a zero-order Sugeno fuzzy model is functionally equivalent to a Radial Bases Function Network under certain minor constraint. The Sugeno fuzzy model is also known as the TSK (Takagi, Sugeno and Kang) model in developing a systematic approach to generate fuzzy rules from a given input–output data set (Berenji & Khedkar, 1992; Takagi & Sugeno, 1985).

4.2. Architecture of hybrid learning and adaptive neuro-fuzzy inference system

Fig. 4 shows the architecture of the neural network structure. The computation of MFs parameters (or their adjustment) is facilitated by a gradient vector, which provides a measure of how well the ANFIS is modeled with the input/output data for a given set of parameters. Once the MFs are constituted, any of several optimization routines can be applied in order to adjust the parameters to reduce the error measure (usually defined by the sum of the squared difference between actual and desired outputs). The parameters associated with the MFs will change through the learning process.

The output of the ith node is denoted in layer l as $O_{l,i}$ which is given in Eq. (4).

Layer 1: every node i in this layer is an adaptive node with a node function

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(Q), \quad \textit{for } i = 1,2, \textit{or} \\ O_{1,i} &= \mu_{A_{i-2}}(M), \quad \textit{for } i = 3,4, \end{aligned} \tag{4}$$

where Q, M or F are the input to node i and $A_{1,\ldots,i}$ is the linguistic label such as *'excellent'*, *'average'* or etc. associated with this node. In other words, $O_{l,i}$ is the membership grade of fuzzy set $A_{1,\ldots,i}$ (linguistic labels) and specifies the degree to which the given assessment tool Q (or M) satisfies the quantifier $A_{1,\ldots,i}$. The membership grade of each linguistic value $A_{1,\ldots,i}$ can be parameterized and calculated using Eq. (5).

$$\mu_A(x) = bell(x;a,b,c) = 1 \Bigg/ \left(1 + \left|\frac{x - c_i}{a_i}\right|^{2b}\right) \tag{5}$$

where $\{a_i, b_i, c_i\}$is the parameter set. Parameters c_i and a_i can be adjusted to vary the center and width of the MF, and then b_i is used to control the slopes at the crossover points with the next MF. As the values of these parameters changes, the bell-shaped function varies accordingly, thus exhibiting various forms of MFs for fuzzy linguistic set $A_{1,\ldots,i}$. Parameters in this layer are referred to as premise parameters (Jang et al., 1997).

Layer 2: every node in this layer is a fixed node and the output of which is the product of all the incoming signals that are presented by Eq. (6).

$$O_{2,i} = w_i = \mu_{A_i}(Q)\mu_{A-2}(M)\ldots\mu_{A-n}(F), \quad i = 1,2\ldots,n \tag{6}$$

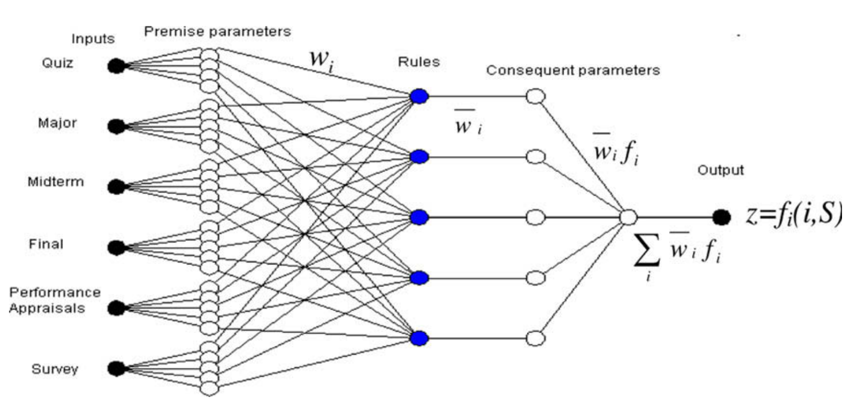

Fig. 4. ANFIS architecture for a six input single-output Sugeno fuzzy model.

Each node output represents the firing strength of a rule. In general, any other T-norm operators that performs fuzzy AND can be used as to node function in this layer.

Layer 3: Every node in this layer is fixed. The *i*th node calculates the ratio of the *i*th rule's firing strength to the sum of all rules' firing strengths using Eq. (7).

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \tag{7}$$

Outputs of this layer are called normalized firing strengths.

Layer 4: Every node *i* is an adaptive node with a node function $O_{4,}i$ as in Eq. (8).

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i(p_i Q + q_i M + \cdots + m_i F + r_i) \tag{8}$$

where $\overline{w}_i$ is a normalized firing strength from Layer 3 and $\{p_i, q_i, \ldots, m_i, r_i\}$are the parameter sets. Parameters in this layer are referred to as consequent parameters.

Layer 5: the single node in this layer is a fixed node labeled Σ, which computes the overall output as the summation of all incoming signals as calculated by Eq. (9).

$$\text{Overall output } z = O_{5,1} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{9}$$

From the ANFIS architecture shown in Fig. 4, it is observed that when the values of the premise parameters are fixed, the overall output can be expressed as a linear combination of the consequent parameters. In symbols, the output *z* can be rewritten as in Eq. (10).

$$\begin{aligned} z &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 + \cdots + \frac{w_n}{w_{n-1} + w_n} f_n \\ z &= \overline{w}_1(p_1 Q + q_1 M + \cdots + m_1 F + r_1) \\ &\quad + \cdots + \overline{w}_n(p_n Q + q_n M + \cdots + m_n F + r_n) \\ z &= (\overline{w}_1 Q)p_1 + (\overline{w}_1 M)q_1 + \cdots + (\overline{w}_1 F)m_1 + (\overline{w}_1)r_1 \\ &\quad + \cdots + (\overline{w}_n Q)p_n + (\overline{w}_n M)q_n + \cdots + (\overline{w}_n F)m_n + (\overline{w}_n)r_n \end{aligned} \tag{10}$$

which is linear in consequent parameters, p_1, p_2, q_1, q_2, r_1, and r_2. From the observations (past student achievements), one may have *S* as the set of total parameters with S_1 and S_2 being the sets of premise and consequent parameters respectively. In this study, the ANFIS contains a total of 95 fitting parameters, of which 60 are premise (nonlinear) parameters and 35 are consequent (linear) parameters. The size of input–output data set is large enough to lay down the ANFIS model and to fine tune the membership functions. A total of 114 training data and 114 checking data have been uniformly sampled from the input ranges and used. The training error is the difference between the training data output value, and the output of the fuzzy inference system corresponding to the same training data input value.

The adaptive network under consideration has only one overall output that is presented by $z = O = f_i(i, S)$ where *i* is the vector of input variables, and *f* is the overall function implemented by the adaptive network Since *f* is linear, one can identify linear parameters by using the linear least-squares method. This approach leads to a hybrid learning rule which combines steepest descent and least-squares estimators for the identification of parameters.

In the least-squares problem, the output of linear model *z* is given by the linearly parameterized expression in Eq. (11).

$$z = \theta_1 f_1(u) + \theta_2 f_2(u) + \cdots + \theta_n f_n(u) \tag{11}$$

where *u* = [*Quiz (Q), Major (M), Midterm (MD), Final (F), Performance Appraisals (P), Survey (S)*] is the model's input vector, $f_1, \ldots, f_2$ are known functions of *u*, and $\theta_1, \ldots, \theta_n$ are unknown MFs parameters to be estimated. The training data set that is needed to identify the unknown parameters θ_i is usually obtained through experiments. This data set represents desired input–output pairs of target system to be modeled. In this study, it was obtained from the past achievement of students and used to develop the ANFIS model. Eq. (11) can be rewritten in a concise form using the matrix notation: $A\theta = z$, where *A* is called design matrix of $f_1, \ldots, f_2$ known functions. θ is an *nx1* unknown parameter vector, and *z* is an *mx1* output vector. The *i*th row of the data matrix [*A*, *z*], denoted by

$[a_i^T, z_i]$ in Eq. (12), is related to the *i*th input–output data pair $(u_i; z_i)$ as the *i*th data pair of the training data set.

$$a_i^T = [f_1(u_i), \ldots, f_n(u_i)] \tag{12}$$

In order to identify uniquely the unknown parameter θ, it is necessary that $m \geqslant n$. However, usually *m* is greater than *n*, indicating the availability of more data pairs than fitting parameters. In this case, an exact solution satisfying all the *m* equations is not always possible, since the data might be contaminated by noise, or the model might not be appropriate for describing the target system. Thus, Eq. (13), $A\theta = z$ should be modified by incorporating an error vector *e* to account the random modeling error, as follows.

$$A\theta + \mathrm{e} = \mathrm{z} \tag{13}$$

The standard back-propagation algorithm uses the steepest descent algorithm to minimize the mean squared error function defined by Eq. (14).

$$E(\theta) = \sum_{i=1}^{m} (z_i - a_i^T \theta)^2 = e^T e = (z - A\theta)^T (z - A\theta) \tag{14}$$

where $e = z - A\theta$ is the error vector produced by a specific choice of θ. In Eq. (14) the squared error is minimized and is called the least-squares estimator (LSE). The least-squares polynomial always fit the data perfectly, but it is not robust; a small amount of noise in the data set could change the whole MFs curves dramatically and make them untrustworthy (Cordon et al., 2004; Jang et al., 1997). To eliminate this problem, another input–output data set, called the validating data set is used. This data set verifies the capability of the resulting ANFIS model and thus provides an unbiased index for selecting the best model and its parameters. A part of training and checking data set and their relative errors are given in Table 1. Fig. 5 presents the relative error of checking and training data found in this study. The average checking error is equal to 0.00105084 and the average root mean square error ($RMSE_{chk}$) for checking data is 0.0324 after 200 epochs. For the same number of epochs, the average training error is 0.000591987 and $RMSE_{trn}$ for training data set is calculated 0.02433. It is unusual to observe that the desired $RMSE_{trn} > RMSE_{chk}$ during the learning process, as is the case here. The role of training and checking data set is changed, the usual situation $RMSE_{trn} < RMSE_{chk}$ is achieved during the learning process. Since both RMSEs are both very small, the ANFIS developed captures the essential components of the underlying dynamics.

4.3. Model validation using checking and testing data set

Model validation is the process by which the input vectors from input–output data sets, are presented to the trained FIS model, to see how well the ANFIS model predicts the corresponding data set of output values. When checking and training data are presented to ANFIS, the selected FIS model is expected to have parameters associated with the minimum checking data model error. The basic idea behind using a checking data set for model validation is that after a certain point in the training, the model begins over-fitting the training data set. In principle, the model error for the checking data set tends to decrease as the training takes place up to the point that over-fitting begins, and then the model error for the checking data suddenly increases. To eliminate this problem, a large amount of data set was collected, as it is seen in the Fig. 6, 214 data is used for

Table 1
Training and checking data and the errors of ANFIS outputs versus the classical statistical method.

Case	Checking data	Checking ANFIS outputs	Checking error (%)	Training data	Training ANFIS outputs	Training error (%)
1	55.3	54.3089	−0.9911	50.85	50.8491	−0.0009
2	56.7	57.1319	0.4319	55.2	55.2	0
3	76.85	77.7111	0.8611	44.5	44.5	0
4	53.7	50.5747	−3.1253	67.2	67.1997	−0.0003
5	63.25	63.4875	0.2375	69.5	69.5	0
6	78.95	77.3034	−1.6466	63	63.0001	0.0001
7	50.75	51.3072	0.5572	72.95	72.95	0
8	62.8	62.2736	−0.5264	56.2	56.2	0
9	65.35	64.4213	−0.9287	42.05	42.05	0
10	72.15	72.2153	0.0653	73.6	73.6	0
11	54.95	54.5746	−0.3754	59.8	59.8	0
12	75.5	72.6474	−2.8526	46.5	46.5	0
13	78.45	80.617	2.167	55.65	55.6499	−1E-04
14	66.6	67.7514	1.1514	69.9	69.9	0
15	68.5	69.3002	0.8002	64.25	64.2498	−0.0002

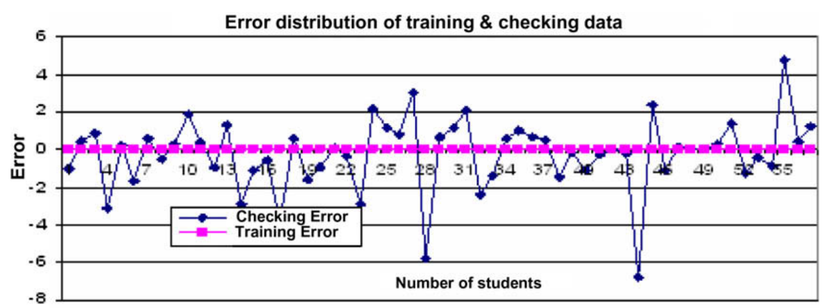

Fig. 5. The error of each observation for checking and training data.

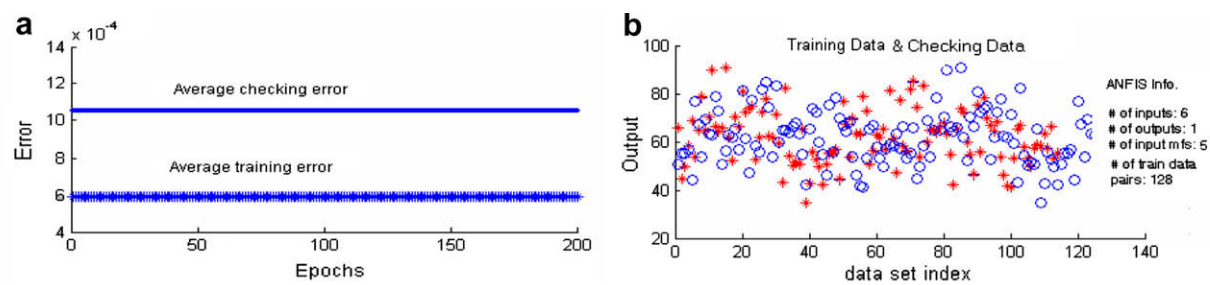

Fig. 6. Average error obtained (a) by training and checking data set (b) to develop the optimal model.

identification each input–output parameter and 100 is used for checking the validation of model. This data set contains all the necessary representative features of the assessment tools.

Since RMSEs are both very small, we conjecture that the ANFIS has captured the essential components of underlying dynamics and the training data contain the effects of the initial conditions which might not be easily accounted for by the essential components identified by ANFIS. The crisp results of fuzzy model that are representative for the SAP is given in Table 1. The SAP assessment usually involves linguistic terms such as *unsatisfactory, average, good, very good, excellent, or bad, very bad, etc.*, which absorbs a substantial amount of fuzziness. The outcomes of the statistical approaches of student academic performance for different input parameters (assessment tools) are calculated and presented in Table 2.

5. Discussion and Conclusions

In this paper, we focused on development of a data-driven ANFIS model using a real dataset obtained from students' achievements in the engineering economy course. The studied ANFIS is a soft computing approach and a general feed-forward multilayer neural network for fuzzy modeling and decision systems. A difficulty facing various existing neuro-fuzzy hybrid learning methods is that the learning is supervised and thus requires training information on the subject domain. In this approach, the fuzzy logic components are directly integrated in the neural networks. The input and output nodes represent the input states and control signals, respectively, and in the hidden layers there are nodes that code membership functions and rules. The learning algorithm used to build rule nodes and training the membership functions is based on the back-propagation algorithm. The ANFIS possesses Sugeno-type reasoning mechanism which is depicted in Fig. 2. The corresponding equivalent ANFIS architecture is shown in Fig. 4. The ANFIS can produce outcomes of the SAP, and compare them with the outcomes of the statistical method. In Table 1, a part of training and checking data set is given. The errors of the ANNs back-propagation method achieved is clearly depicting that the model is successful for the SAP applications. A comparison of the outcomes of the ANFIS model and the statistical method is given in Table 2

Table 2
Outcomes of the SAP for statistical and ANFIS model; their linguistic values based on different inputs.

Case	Q	M	MD	F	P	S	SAP statistical	SAP ANFIS	Linguistic values
1	75	67	46	50	75	50	56.75	56.74	Average
2	40	30	38	40	67	40	40.8	40.81	Unsatisfactory
3	30	45	32	60	75	50	50.15	50.15	Unsatisfactory
4	70	87	75	90	95	72	84.15	84.14	Very well
5	66	72	65	95	80	62	79.5	79.5	Very well
6	40	38	50	67	75	61	57.05	57.06	Average
8	55	75	71	85	90	50	76.45	76.45	Very well
9	70	95	85	65	80	73	75.9	75.9	Very well
10	73	85	75	100	95	61	87.6	87.61	Excellent
11	5	50	58	75	90	50	61.1	61.1	Average
12	100	88	100	94	45	56	88.1	88.1	Excellent
13	33	70	63	65	80	65	63.65	63.66	Average
14	50	74	56	85	85	67	73.15	73.15	Good
15	57	53	56	80	70	54	66.55	66.55	Average

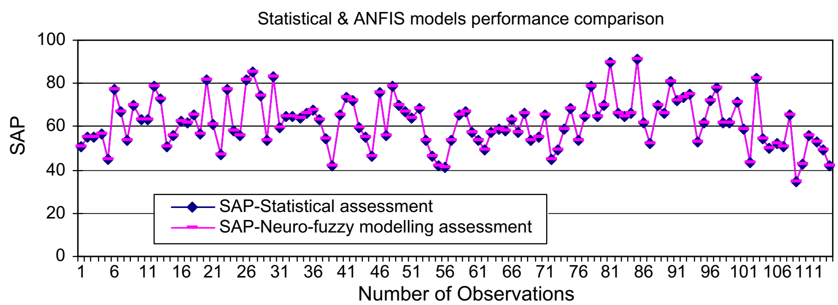

Fig. 7. The outcomes of SAP versus Statistical and ANFIS for different observations.

Table 3
ANFIS SAP outcomes with subclustering and grid partitioning methods, their error versus statistical SAP outcomes.

Q	M	MD	P	S	F	Statistical SAP outcomes	ANFIS SAP outcomes, subclustering	Error	ANFIS SAP outcomes, grid partitioning	Error
60	48	25	60	80	60	50.85	50.8502	−0.0002	51	−0.15
50	67	46	50	75	50	55.3	55.3001	−0.0001	56	−0.7
60	76	57	60	75	50	55.2	55.2001	−1E-04	55.9999	−0.7999
85	70	46	85	75	53	56.7	56.6999	0.0001	57	−0.3
60	45	32	60	75	50	44.5	44.5	0	41	3.5
90	87	75	90	95	72	76.85	76.8496	0.0004	77.0001	−0.1501
95	72	65	95	80	62	67.2	67.2	0	68	−0.8
67	38	50	67	75	61	53.7	53.7	0	49.9999	3.7001
85	73	60	85	80	72	69.5	69.5	0	69.9999	−0.4999
85	75	71	85	90	50	63.25	63.25	0	65.4999	−2.2499
67	66	70	67	65	70	63	63	0	63	0

and illustrated in Fig. 7. There is a clear indication that the outcomes of SAP for the statistical method and the ANFIS model closely agree for the same set of input variables.

Table 3 provides a comparison of outcomes obtained using ANFIS subclustering and grid partitioning methods with the statistical methods for the same set of input variables. Fig. 8 depicts the scatter plot of 112 principal component outcomes of the ANFIS model. Normally, each outcome represents the achievement of a student falling in a category and has a numerical value representing it. The outcomes in Figs. 7 and 8 show that the ANFIS model is able to produce the same outputs comparing with the statistical method. This assessment component would allow more flexibility to make judgment on a single or a group of student performances and would be more understandable by parents as well as the educators and trainers for further evaluations of past achievements and future directions. The student's achievements are expressed by quality indicators of learning performance in five linguistic terms starting with A_1 *(unsatisfactory)*, A_2 *(average)*, A_3 *(good)* and A_4 *(very well)*, ending with A_5 *(excellent)*. Normally, these expressions are used in our daily life to grade the quality of products or services. For instance, students who scored 80 and above might be considered as a group and the performance of them might be expressed by a linguistic value A_5 *(excellent)*. This linguistic value is a score showing that a student in this batch has achieved an excellent performance. As an example, if a student scores 85 out of 100, will partially belong to both 'A_4 *(very well)*' set with a degree of membership of 0.8 ($\mu_A(85) = 0.8$) and 'A_5 *(excellent)*' set with a degree of membership 0.55 ($\mu_A(85) = 0.55$). Similarly, students within the 50–60 score represent a group and the performance of these students might be expressed linguistically A_2 *(average)*. A student scored 55 out of 100 will be a member of set A_2 *(average)* with a degree of membership of 0.9 ($\mu_A(55) = 0.9$) and would be an average student in his/her academic performance. The same student would also belong to set A_1 *(unsatisfactory)* also with a degree of membership

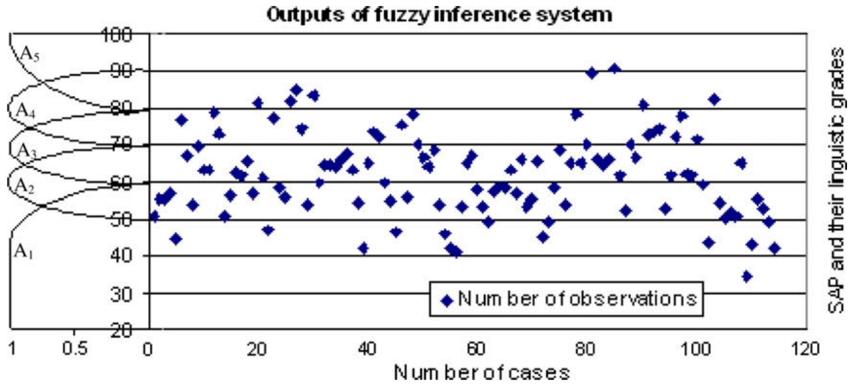

Fig. 8. The outcomes of ANFIS model and their linguistic values.

($\mu_A(55) = 0.5$). This is the flexibility of fuzzy linguistic variables and their labels.

Assessment of students' achievements in the previous year to evaluate the learning process and performance of students might be more meaningful than referring the marks individually. The previous and current student's performance data are kept in a computer system that may be an excellent source to achieve satisfactory results in evaluation of the students' performance as a group, their learning level and evaluate the quality of education. Fig. 7 illustrates students' achievements, academic performances, linguistic grades granted. Table 2 also gives the outcomes of the SAP for statistical and neuro-fuzzy modeling approaches with their linguistic values.

In conclusion, several reasons exist to assess the performance of students. Performance assessment gives feedback on the effectiveness of teaching, the extent to which the course aims have been achieved, and information on the effectiveness of learning. Different evaluation methods have been used for primary, secondary and tertiary education. A variety of them also exist in different countries around the world. In mostly used statistical methods, different scores of each assessment tool are added up based on determined weights to obtain a single score for an individual student's performance. However, evaluators usually lack a formal reasoning mechanism to support the inference and typically marks are decided according to given marking schemes, experiences, sensitivities and standards. Thus, marks assigned by an evaluator are only approximations. Conversely, the academic performance evaluation involves the measurement of ability, competence and skills which are fuzzy concepts and can be approximated by fuzzy linguistic terms. Eventually, linguistic terms can be awarded to a single student's achievement as well as a group of students who had already taken a course.

References

Berenji, H. R., & Khedkar, P. (1992). Learning and tuning fuzzy logic controllers through reinforcements. *IEEE Transactions on Neural Networks, 3*, 724–740.

Chowdhury, M. M., & Li, Y. (1998). Learning fuzzy control by evolutionary and advantage reinforcements. *International Journal of Intelligent Systems, 13*, 949–974.

Cordon, O., Gomide, F., Herrera, F., Hofmann, F., & Magdalena, L. (2004). Ten years of genetic fuzzy systems: Current framework and new trends. *Fuzzy Sets and Systems, 141*, 5–31.

Deniz, D. Z., & Ersan, I. (2002). An academic decision-support system based on academic performance evaluation for student and program assessment. *International Journal of Engineering Education, 18*(2), 236–244.

Ebel, R. L., & Prade, H. (1991). *Essentials of educational measurements* (5th ed.). New Jersey: Prentice-Hall.

Figueiredo, M., & Gomide, F. (1999). Design of fuzzy systems using neurofuzzy networks. *IEEE Transactions on Neural Networks, 10*(4), 815–827.

Ishibuchi, H., Nii, M., & Turksen, I. B., (1998). Bidirectional bridge between neural networks and linguistic knowledge: Linguistic rule extraction and learning from linguistic rules. In *Proceedings of the IEEE International Conference Fuzzy Systems* (p 1112–1117).

Ishibuchi, H., Nakashima, T., & Murata, T. (1999). Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. *IEEE Transactions on Systems Man and Cybernetics, 29*, 601–618.

Jang, J. S. R., & Sun, C. T. (1993). Functional equivalence between radial basis function networks and fuzzy inference systems. *IEEE Transactions on Neural Networks, 4*, 156–159.

Jang, J., & Sun, C. (1995). Neuro-fuzzy modeling and control. *Proceedings of IEEE, 83*, 378–406.

Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). *Neuro-fuzzy and soft computing*. New Jersey: Prentice-Hall.

Karagözoğlu, B., (2007). Description of capstone courses for biomedical engineering program at King Abdulaziz University. *Proceedings of the second international conference on engineering education & training, Kuwait*.

Karagözoğlu, B., & Türkmen, N., (2007). A software tool to facilitate design, assessment and evaluation of courses in an educational system. In *Proceedings of the fourth IEEEGCC conference, Bahrain*.

Law, C. K. (1995). Using fuzzy numbers in educational grading system. *Fuzzy Sets and Systems, 83*, 311–323.

Lin, C., & Lee, C. (1991). Neural-network based fuzzy logic control and decision system. *IEEE Transactions on Computer, 40*, 1320–1336.

Lin, C., & Lu, Y. (1996). A neural fuzzy system with fuzzy supervised learning. *IEEE Transactions on Systems Man and Cybernetics, Part B, 26*, 744–763.

Lopes, A. L. M., Lanzer, E. A., & Barcia, R. M., (1997). Fuzzy cross-evaluation of the performance of academic departments within a university. In *Proceedings of Canadian Institute Research and Planning Association conference*, 19–21, Toronto, Canada.

Ma, J., & Zhou, D. (2000). Fuzzy set approach to the assessment of student-centered learning. *IEEE Transactions on Education, 43*(2), 237–241.

Magdalena, L., & Monasterio, F. (2000). A fuzzy logic controller with learning through the evolution of its knowledge base. *International Journal of Approximate Reasoning, 16*, 335–358.

Mamdani, E. H. (1997). Application of fuzzy logic to approximate reasoning using linguistic systems. *International Journal of Man–Machine Study, 7*(1), 1–13.

Mendel, J. M. (2001). *Uncertain rule-based fuzzy logic systems: Introduction and new directions*. New Jersey: Prentice-Hall PTR.

Mitra, S., & Hayashi, Y. (2000). Neuro-fuzzy rule generation: Survey in soft computing framework. *IEEE Transactions on Neural Networks, 11*(3), 748–768.

Nie, J. (1995). Constructing fuzzy model by self organizing counterpropagation network. *IEEE Transactions on Systems Man and Cybernetics, 25*, 963–970.

Pal, K. S., & Mitra, S. (1999). *Neuro-fuzzy pattern recognition: Methods in soft computing*. New York: Wiley.

Papadakis, S. E., & Theocharis, J. B. (2002). A GA-based fuzzy modeling approach for generating TSK models. *Fuzzy Sets and Systems, 131*(1), 121–152.

Pedrycz, W. (1993). *Fuzzy control and fuzzy systems*. New York: Wiley.

Ratcliff, K. A., Arkin, R. M., & Dove, M. K., (1997). A method for assessing undergraduate general Education. In The 37th AIR Annual Forum, Florida, May 18–21.

Schmucker, K. J. (1989). *Fuzzy sets, natural language computations and risk analysis*. Maryland: Computer Science Press.

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems Man and Cybernetics, SMC-15*, 116–132.

Taylan, O. (2006). Neural and fuzzy model performance evaluation of a dynamic production system. *International Journal of Production Research, 44*(6), 1093–1105.

Taylan, O., & Karagözoğlu, B. (2007). A fuzzy rule-based modeling approach for student academic performance evaluation. In *Proceedings of the second international conference on engineering education & training, Kuwait, 9–11 April.*

Zadeh, L. A. (1972). A fuzzy set theoretic interpretation of the linguistic hedges. *Journal of Cybernetics, 2*, 4–34.