Adaptive Neuro-Fuzzy Inference System in Predicting the Success of Student's in a Particular Course

Sümeyye Kaynak, Hayrettin Evirgen, and Baran Kaynak

Abstract—**Students are required to choose a certain number of elective courses for next semester. One of the factors influencing the course selection of the students is their grade in the course. The grades that students will get from the course at the end of the semester depend on many factors and includes uncertainties. The student's gpa, gender and grades of related courses are the parameters that can be obtained to eliminate this uncertainty. However, the parameters such as student's psychological state, etc. are difficult to obtain, and will eventually have a large impact on his/her grade at the end of the semester. Estimate implementation depends on many factors and the degree that these factors affect the estimation is uncertain. In this study, using adaptive neuro-fuzzy model, the estimation of the score range that the student will complete the courses s/he requests is analysed and an average result of ±0.36 points is found.**

Index Terms—**ANFIS, decision support system, education.**

I. Introduction

Students have to take a certain number of compulsory and a certain number of elective courses each semester. Students are required to be successful in these courses in order to graduate. The decisions that students make in course selection can directly affect their future [1]. While choosing courses, students consider a number of factors such as the grade that they will get from the course, obtaining knowledge, the usability of knowledge, instructor's characteristics, course hours, course time and whether they are interested or not. Kardan, Sadeghi, Ghidary and Sani (2013) states in their article that course characteristics, instructor's characteristics, student's workload, course grade, course type, course time, number of time conflicts, final examination time and student demands are important factors in course selection. Even, sometimes grades rather than learning become the primary goal of students. They might need to earn high grades for future admission into advanced programs, applying for a well-paying job, or any other personal reason [2].

Students determine the courses they will select at the beginning of each term. The primary objective of the students is to complete the selected course with a high grade. Therefore, the estimation of the final score of the course to be selected by the students will guide students in course selection. The realization of this estimate depends on multiple parameters. These parameters are not directly relevant to the output and it is not known how much each of these parameters affects the results. The estimation of the final score of the course to be selected by the student is a nonlinear problem.

Adaptive neural fuzzy inference systems are tools used in nonlinear problem solving [3]. ANFIS is used for estimation purposes similarly to neural networks [4]. Further information about ANFIS is given in the next section.

The actual 10-year student dataset of Sakarya University Computer Engineering Department is used for the current study. With this dataset, the success to be achieved by the student from the courses selected is estimated using ANFIS method. ANFIS model has multiple inputs and a single output. The inputs of ANFIS model are the grade point average at the end of the 6th period, the average success of the courses related to the course, and student gender. The output of ANFIS is the success degree of the course selected by the student at the end of the semester.

In ANFIS applications, classifying the whole dataset into training and test data is a common method. Training and test data are selected randomly from the whole training set. It is expected that training set have an example space including all situations. Training data set is used in building the model of ANFIS method, whereas test data set is exploited in evaluating the performance of the model. This practice is a sufficient method for many studies. But, it will not be suitable for some circumstances. As Zhang, Hu, Patuwo, Indro (1999) stated in their articles; evaluation in that the characteristics of the test may be very different from those of the training and the estimated classification rate can be very different from the true classification rate particularly when small-size samples are involved [5]. For all these reasons, cross-validation method is used in this application. This method is a model evaluation method. Owing to this method, the performance of the ANFIS model can be more accurately defined. Also the cross-validation analysis yields valuable insights on the reliability of the ANFIS mfodel with respect to sampling variation [5]. Further information about cross validation is given in the Section III.

Due to the fact that this study is a decision support application, it helps students, students' advisors, course instructors, and even the university administration. With the decision support system in our study; consultant can estimate the course achievement of the student, can recommend the course to him/her and can understand whether the course is of great interest to him/her without having to know him/her in

Manuscript received February 10, 2014; revised April 21, 2014.

S. Kaynak is with the Institute of Natural Sciences, Department of Computer and Information Systems Engineering, Esentepe Sakarya Turkey (e-mail: sumeyye@sakarya.edu.tr).

H. Evirgen is with the Faculy of Computer and Information Sciences, Department of Computer Engineering, Esentepe Sakarya Turkey (e-mail: evirgen@sakarya.edu.tr).

B. Kaynak is with the Faculy of Computer and Information Sciences, Department of Information Systems, Esentepe Sakarya Turkey (e-mail: kaynak@sakarya.edu.tr).

DOI: 10.7763/IJCTE.2015.V7.926

depth. Instructor can make changes about the course content, the course flow or can make improvements if necessary.

The number of students per counselor is very high in Sakarya University. This situation makes it difficult to do the consulting services for students. In addition, students in different cities make it difficult to communicate with their advisors. This application will reduce the dependency on the student's advisor in the selection of courses.

The article consists of 5 sections. The rest of this paper is structured as follows: Section II describes an introduction to the theoretical basis of the Adaptive Neuro-Fuzzy systems. Section III describes an introduction to the Cross-Validation procedure. Section IV continues with data collection and preparation of data for application. Section V discusses student achievements. Section VI presents the discussions and conclusions of the research.

II. Explanation of the ANFIS Model

By embedding the fuzzy inference system into the framework of adaptive networks, we obtain the ANFIS architecture [6]. Fuzzy inference systems map the supplied input to the related outcomes using fuzzy logic [7]. Fuzzy inference engine can be implemented using either Mamdani type or Sugeno type [8]. The fundamental difference between Mamdani-type inference systems and Sugeno-type inference systems is the way the crisp output produced [6], [9], [10]. The technique of fuzzy output clarification is used in Mamdani type inference systems, while the weighted average is used in Sugeno-type inference systems [9].

ANFIS is a Sugeno type inference system [7]. It is a hybrid method using the parallel calculation and learning ability and fuzzy logic inference feature of ANN [11]. ANFIS architecture consists of 5 layers [12], [13] as shown in Fig. 1. Each layer contains several nodes described by the node function [13].

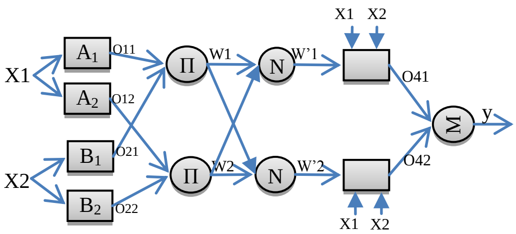

Fig. 1. ANFIS architecture with two inputs.

A. Layer 1

Each node in the first layer of ANFIS architecture processes the inputs coming by using node functions. The input values of the node are $X1$ and $X2$. A_i and B_i represent fuzzy sets. The output of this layer is the inputs' degree of membership in fuzzy sets. Fuzzy sets have membership functions. Inputs are members of fuzzy sets as much as the output values of the membership functions. There are many membership functions such as triangular mf, generalized bell mf and Gaussian mf [11]. In our study, generalized bell mf (Gbellmf) was used. The output of each node is defined by (1).

$$\begin{aligned} O_{1i} &= \mu\mu_i(x1)\ for\ i=1,2 \\ O_{2i} &= \mu\mu_i(x2)\ for\ i=1,2 \end{aligned} \tag{1}$$

For example;

$$\mu A_i(x1) = \frac{1}{\left(1+\left|\frac{x1\text{-}c_i}{a_i}\right|^{2b}\right)} \tag{2}$$

a_i, b_i and c_i in the formula (2) are the parameters of this layer. The parameters in this layer are called premise parameters. The output of the node is also subject to change together with the changes in the values of these parameters.

B. Layer 2

This is a layer of rules. The outputs of the first layer constitute the inputs of this one. $W1$ value is the multiplication of the output of A_1 node and that of B_1 node (3). Likewise $W2$ value is the multiplication of the output of A_2 node and that of B_2 node (4). $W1$ and $W2$ values show the ignition strength of the rule.

$$W1 = O11 * O21 = \mu A_1(x1) * \mu B_1(x2) \tag{3}$$

$$W2 = O12 * O22 = \mu A_2(x1) * \mu B_2(x2) \tag{4}$$

C. Layer 3

This is a layer of normalization. All the outputs of the nodes in the layer of rules are used as input. The proportion of the ignition power of the node i in the layer of rules to the sum of the ignition power of all the nodes gives the normalized ignition rate of node I (5).

$$W'i = Wi + / (W1 + W2),\ i = 1,2. \tag{5}$$

D. Layer 4

This is the clarification layer. Node i in this layer compute the contribution of i-th rule toward the overall output [3].

$$W'iF_i = W'i(p_i X1 + q_i X2 + r_i) \tag{6}$$

The parameters p_i, q_i, r_i where in the formula (6) are consequent parameters.

E. Layer 5

There is one node in this layer. This node sums the output values of each node in the layer 4. This summation is the output value of the ANFIS system.

ANFIS system performance can be measured using various statistical methods. In the application, the root mean squared error (RMSE) is used to measure the ANFIS system performance. RMSE, is the square root of the sum of the squares of the difference between the actual system output and the model output. Our goal is to minimize this error value.

III. Explanation of the Cross-Validation Procedure

Cross-validation is a model evaluation method that provides an indication of how well the learner will do when it is asked to make new predictions for data it has not already seen [14]. Several methods for implementing the cross-validation theory were proposed in the literature; however, the essence of all these methods is similar[15]. Cross-validation method, basically, classifies the whole data set into 2 groups as training and test data. The model is trained with training data and performance of the trained model is measured with the test data.

K-fold cross validation is a kind of cross validation. It bears the basic idea of cross validation. One part of the data set is used for training and the other part is benefited for test. It separates the original data set into k sets. Each set separated from the original one is called as "fold". Training and test data sets are generated through various combinations of these folds. The algorithm of the K-Fold cross validation is represented through an example.

1) Suppose there are 100 data for application.
2) Divide the whole data set into 5 subsets. K value =5
3) Each subset (fold) includes 20 data.
4) The entire data split in two group as training set and testing set. Let us divide the dataset 80 for training, 20 for test. So, 4 folds for training and 1 fold for testing are kept.
5) Among 5 fold, 4 fold are selected randomly for training and the remaining one for testing.
6) With these 4 fold, the performance of the trained model is evaluated by the 1 fold divided for testing.
7) The average testing error is calculated.
8) 4 out of 5 folds are reselected for training. The remaining fold is used for testing. All the possible combinations are shown in Table I. Each possible situation is an iteration.

TABLE I: K-Fold Cross Validation Example

Iteration	Fold for training	Fold for testing
1	1,2,3,4	5
2	1,2,3,5	4
3	1,2,4,5	3
4	1,3,4,5	2
5	2,3,4,5	1

9) 7. and 8. steps are repeated for each iteration.
10) The average error values extracted from each iteration are summed and divided by the number of iterations. This is the value of the average error of the application.

IV. Data Collection and Preparation of Data for the Application

Having begun education in 1994, Sakarya University Computer and Information Sciences Faculty of Computer Engineering department graduated 1214 students until 2013. The students of Sakarya University Computer Engineering have elective courses in 7th and 8th semesters. Elective courses are divided as; technical, social and university common electives. The application is conducted on data mining courses, which is a technical elective.

Data mining course, which is a 7th semester technical elective, has been given by Sakarya University Department of Computer Engineering since 2005. Since 2005, a total of 447 students have selected the course. Data mining course is among the most preferred courses by Sakarya University students.

The GPA of the student at the end of the 6th period, the grades received from the compulsory courses that may be related to the data mining course (probability and statistics, numerical analysis, discrete practical structures) and GPA of these courses, and the student's gender information are used in the application. This information is taken from the student information database of Sakarya University. The information retrieved from the database is transferred to Excel in C #.

It can be seen in the data set taken that some of the students did not enrol in some or none of the compulsory courses to be used as input; some did not select the elective data mining course; and some were exempted from these classes or competent at them. The student data having such features has been removed from the data set.

It is a problem that students with missing courses of the related ones have no grades. To overcome this problem, the student's grade averages of other related courses are used. Alternatively, a student's periodical grade point average is used. However, both options mislead the result. So they were removed from the data set. Students who are exempt from the related courses or well enough were removed from the data set as their grades cannot be represented with any numeric value. Following all of these operations, the number of students has dropped from 2128 to 240.

The genders of students are represented with the values 1 and 2. Value 1 represents female students and value 2 represents male students. Periodic grade point average of the student is evaluated out of 4. At Sakarya University, course grades of students are represented by the letters: AA, BA, BB, CB, CC, DC, DD and FF. AA letter indicates the best grade, DD indicates the worst grade and FF indicates that the student fails the course. In the application, the values between the letters AA and FF are represented in the scale from 1 to 8.

V. Student Achievement Prediction Using ANFIS Model

In the application, the estimation of students' grade levels, at the end of the term, in the courses s/he selects is analysed using ANFIS model sugeno-type fuzzy inference system with cross validation. Matlab 2012 platform has been utilized in the implementation of ANFIS model, and the tools MSSQL 2012 and Excel 2013 have been used in data preparation.

A membership function is assigned for each input in ANFIS model and the number of membership functions can be different for each input. Depending on the characteristics of the input set, different membership functions can be selected. Each input is made fuzzy with the membership functions assigned and each of the fuzzy inputs is processed in the layers of ANFIS model. These operations are performed by the rules. As a result of the rules, an output is created. There are two different types of output membership functions in order to create this output. First-order output membership function is a linear function of the type *ax* + *by* + *c*. In the zero-order output membership function, coefficients aand *b* are equal to 0. Thus, it is equal to the constant *c* and this function is called a constant output membership function.

In the application; model performances are evaluated by testing the options that constitute the ANFIS model in different ways. Two different datasets were used as ANFIS model inputs. The inputs of the first dataset are student's gender, the grade point average at the end of the 6th semester and the year-end grade points of each course related to data mining course. In the second set of data entry, the grade point average of them is given as an ANFIS input, instead of giving the year-end grade points of each course related to the data mining course separately. The input of the second data set showed the best performance.2 membership functions for the gender input to ANFIS, 4 for the GPA at the end of the 6th semester, and 8 for the GPA of the related courses were assigned. A Gbellmf membership function was selected as input membership function and a linear membership function as an output membership function. The hybrid learning rule was selected as a learning rule, which combines the gradient method and the least squares estimator (LSE) for identification of parameters [6]. Epoch number was selected as 50.

In this study, a sixfold cross-validation is used. All the examples are divided into 6 equal parts. Each of these parts is called "Fold". 5 random fold for training dataset and 1 random test data set are selected. The selection can be made by 6 different ways. Each of these selection is an iteration of the application. ANFIS model created with training data. The ANFIS model has created 64 rules (Fig. 6). An output is produced with the guidelines established in ANFIS model. The result is the output of the ANFIS model.

At first iteration, the first 5 fold are used for training of the model. 6^{th} fold is exploited for testing. Model was checked through the testing data after training. Checking (testing) the model via testing data represents the average testing error of the model. The average testing error at this iteration is 0.46. The average testing errors calculated from the other iterations are displayed in Fig. 2.

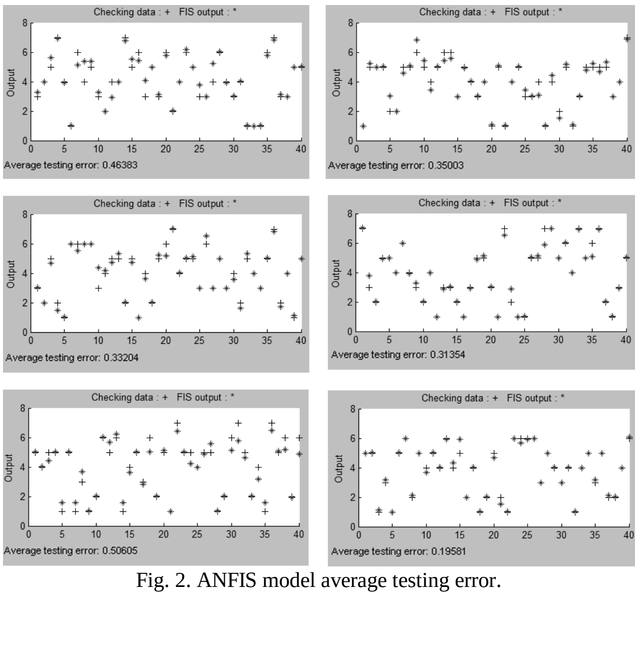

Fig. 2. ANFIS model average testing error.

As can be shown in Fig. 3, 5^{th} fold at the 2^{nd} iteration, 4^{th} fold at the 3^{rd}iteration, 3^{rd} fold at the 4^{th} iteration, 2^{nd} fold at the 5^{th} iteration and, finally, 1^{st} fold at the last iteration were used for testing. The folds exploited for training in Table II, the folds for testing, the average training error of the model and the average testing error values of the model are represented in a way to assign each row to an iteration. In Fig. 4, the graphs demonstrating the average error value of the model for each epoch during the training phase of the ANFIS are presented. Each of these graphs symbolize the iterations of K-fold cross validation procedure.

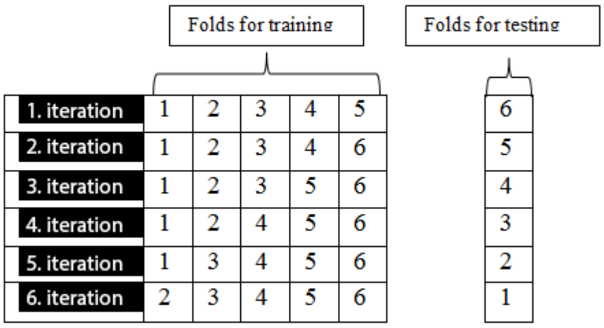

Fig. 3. K-fold cross validation iterations.

TABLE II: K-FOLD CROSS VALIDATION ERROR RESULTS

Fold for training	Fold for testing	Average training error	Average testing error
1,2,3,4,5	6	0.39	0.46
1,2,3,4,6	5	0.35	0.35
1,2,3,5,6	4	0.29	0.33
1,2,4,5,6	3	0.32	0.31
1,3,4,5,6	2	0.31	0.50
2,3,4,5,6	1	0.29	0.19

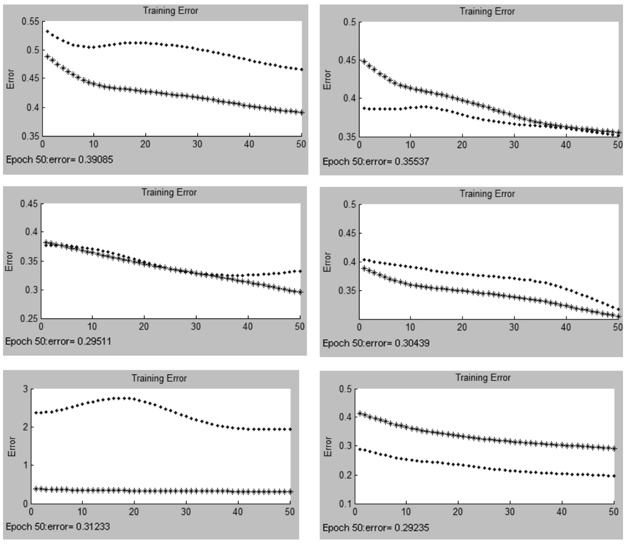

Fig. 4. Average error obtained by training and checking data.

Output values between 1 and 8 were produced using the inputs and outputs of the model in training. The outputs are represented with y-axis in the graph in Fig. 5. However, it can be seen in Fig. 5 that the outcomes estimated by the model (*) do not coincide exactly with the actual outputs. It can be interpreted that the estimated value deviates at an average value of 0.36 with RMSE, which is the square root of the sum of the squares of the distances to the target value. This error value indicates that the model did not do memorization it but made an approach.

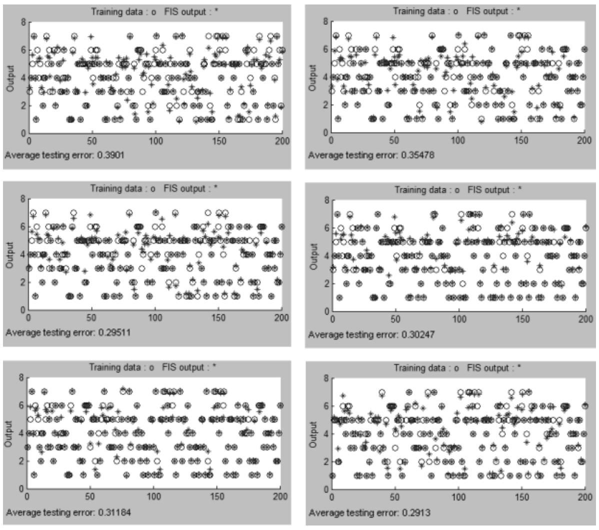

Fig. 5. ANFIS model average training error.

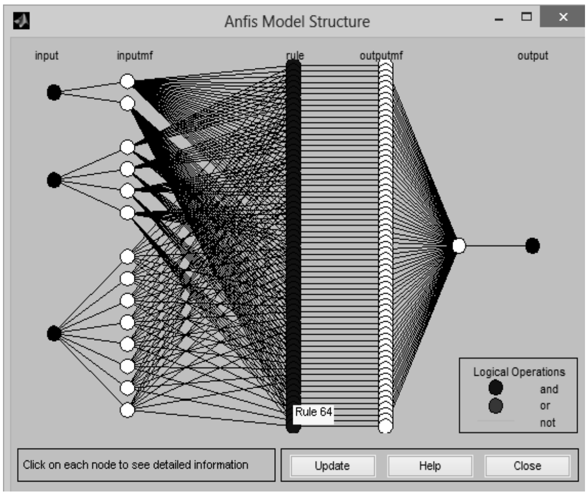

Fig. 6. ANFIS model structure.

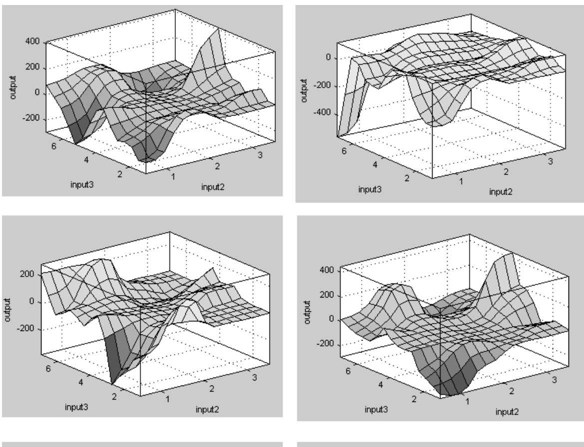

Fig. 7. Three-dimensional representation of all possible values of input2 (GPA in the 6th semester) and input3 (the average grade of the courses related with data mining course) of the ANFIS model and ANFIS output value corresponding to these values.

The number of membership function that is pre-defined for each of ANFIS model inputs directly affects the number of rules in the rule layer. It is observed that 64 rules arise as a result of pairing the membership function of each input seen in Fig. 6 with the membership functions of other inputs. It becomes possible through this screen that ANFIS model can be edited by an expert. An expert can edit the relevant rule by. More uniform results can be obtained with this process in the applications that are well known by specialists. In this study, there was no need to edit any rules.

When Fig. 7 is examined, some of the output values are seen to be out of range. As the reason for this is the absence of an example representing the *x* and *y* values used during training. For example; there does not exist data on a student with a GPA of 4 in the 6th semester and with related courses having an average grade of 7 (DD) in the training data set. The ANFIS model was not trained for such a situation. Therefore, ANFIS model cannot produce an accurate output for these input values.

VI. Discussion and Conclusions

In this application, the estimate of the score that the student will get prior to choosing electives was examined with the ANFIS model. Data mining course was chosen as a sample in this application.

With the application performed, the estimation of students' grades was realized in accordance with the examples used to train the model. The RMSE value of the estimate between 1 and 8 Values was 0.36. The estimation is conducted according to the related courses of the students who want to select it and his/her average semester score, students' gender, and an average result of 0.36 points up or down is found.

A student's grade is very difficult to estimate exactly. Because, the parameters such as instructor's relationship with student, his/her psychological state, his/her interest in the course, his/her health, his/her performance at the time of exam make it difficult to make an estimation. Therefore, showing students the score which is estimated at regular intervals will give more accurate results. In the application, an upper and a lower degree of success are taken as the range. The purpose here is to provide the student with a counselling service along with a decision support system. By presenting the grade range that they will get at the end of the year to the students, benefits are provided to both the student and the advisor.

The calculated ANFIS can be a decision support system for the students who want to select the course calculated by utilizing it each semester. In order to operate the ANFIS model calculated on MATLAB, the model should be shared with students and each student should perform his/her estimation through MATLAB. However, this solution is impractical. Students are requested to provide the necessary inputs through web based programming, which is a more practical way, and the estimate can be calculated with the ANFIS model generated from a remote server. Calculated outputs can be presented to students via a web-based program.

As the ANFIS model creates the rules depending on the inputs; it can give out-of-range results if very different examples from the training inputs are given. The fact that training data set represents the range of values as well as possible gives more efficient results in order to reduce this type of unwanted situations. Filtering extreme cases that can be considered as unusual before being passed to the ANFIS model as inputs has given more accurate results.

REFERENCES

[1] E. Babad, "Students' course selection: Differential considerations for first and last course," *Research in Higher Education*, vol. 42, no. 4, pp. 469–492, 2001.

[2] A. A. Kardan, H. Sadeghi, S. S. Ghidary, and M. R. F. Sani, "Prediction of student course selection in online higher education institutes using neural network," *Comput. Educ.*, vol. 65, pp. 1–11, Jul. 2013.

[3] R. Babuška and H. Verbruggen, "Neuro-fuzzy methods for nonlinear system identification," *Annu. Rev. Control*, vol. 27, no. 1, pp. 73–85, Jan. 2003.

[4] M. Caner, "Mermer Kesme İşleminde Spesifik Enerji Faktörünün ANFIS ve YSA Yöntemleri ile Tahmini. Estimation of Specific Energy Factor in Marble Cutting Process Using," Pamukkale üniversitesi, *Mühendislik Bilimleri Dergisi*, pp. 221–226, 2009.

[5] G. Zhang, M. Y. Hu, B. E. Patuwo, and D. C. Indro, "Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis," *Eur. J. Oper. Res.*, vol. 116, no. 1, pp. 16–32, Jul. 1999.

[6] J. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. Syst. Man. Cybern.*, vol. 23, no. 3, pp. 665–685, 1993.

[7] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst. Man. Cybern.*, vol. SMC-15, no. 1, pp. 116–132, Jan. 1985.

[8] G. Raju and V. M. Sumalatha, "Solving Uncertain Problems using ANFIS," *Int. J. Comput. Appl.*, vol. 29, no. 11, pp. 14–21, 2011.

[9] A. Kaur and A. Kaur, "Comparison of Mamdani-Type and Sugeno-Type Fuzzy Inference Systems for Air Conditioning System," *Int. J. Soft Comput. Eng.*, vol. 2, no. 2, pp. 323–325, 2012.

[10] M. A. Çakiroğlu, E. Erenoğlu, S. Kasap, and E. Yunus, "Çelik Lif Katkılı Betonların Tahribatsız Deney Yöntemleriyle Elde Edilen Basınç Dayanımının ANFIS Metoduyla Tahmini," *SDU Int. J. Technol. Sci.*, vol. 3, no. 3, pp. 14–22, 2011.

[11] Ö. Demirel, A. Kakilli, and M. Tektaş, "ANFIS ve ARMA Modelleri İle Elektrik Enerjisi Yük Tahmini," *Yüksek Lisans Tezi*, vol. 25, no. 3, pp. 601–610, 2009.

[12] A. Gümüş and A. Güneri, "An integrated bullwhip effect measurement system for fuzzy multi-echelon supply chains," *Sigma*, no. 212, pp. 314–324, 2008.

[13] H. Yan, Z. Zou, and H. Wang, "Adaptive neuro fuzzy inference system for classification of water quality status," *J. Environ. Sci.*, vol. 22, no. 12, pp. 1891–1896, Dec. 2010.

[14] A. A. Najah, A. El-Shafie, O. A. Karim, and O. Jaafar, "Water quality prediction model utilizing integrated wavelet-ANFIS model with cross-validation," *Neural Comput. Appl.*, vol. 21, no. 5, pp. 833–841, Nov. 2010.

[15] A. Noureldin, A. El-Shafie, and M. R. Taha, "Optimizing neuro-fuzzy modules for data fusion of vehicular navigation systems using temporal cross-validation," *Eng. Appl. Artif. Intell.*, vol. 20, no. 1, pp. 49–61, Feb. 2007.

Sümeyye Kaynak started to pursue her bachelor's degree in computer engineering at Sakarya University, Turkey in 2008 and completed her bachelor's degree with the third degree. Started to her master's degree in 2012 at Faculty of Computer and Information Sciences, Sakarya University. She has been working in Institute of Natural Sciences at Sakarya University as research assistant since 2012. Interested in artificial neural networks, neuro-fuzzy, genetic algorithms, big data, and data mining.

Hayrettin Evirgen completed his bachelor's degree in physics engineering at Selçuk University, Turkey from 1980 to 1986. He completed his master's degree in business at Sakarya University, Turkey from 1995 to 1997. He completed his Ph.D. in production management at Sakarya University, Turkey from 1998 to 2002. He has been working in computer engineering at Sakarya University as assistant professor. Interested in data mining, and computer software.

Baran Kaynak started to pursue his bachelor's degree in industrial engineering at Sakarya University, Turkey in 2007. And he started to pursue his master's degree in 2011 at Industrial Engineering, Sakarya University. He has been working in the Department of Information Systems Engineering at Sakarya University as academician since 2013. He is interested in object oriented software, neuro-fuzzy, genetic algorithms, job scheduling, big data, and data mining.