

Research Article

Student Consultancy Service: Prediction of Course Grades in Course Selection Phases Using Artificial Intelligence Techniques

Sümeyye KAYNAK^{*1}, Baran KAYNAK², Hayrettin EVİRGEN³

¹ Sakarya University, Faculty of Computer and Information Sciences, Department of Computer Engineering, sumeyye@sakarya.edu.tr

² Sakarya University, Faculty of Computer and Information Sciences, Department of Information Systems, kaynak@sakarya.edu.tr

³ Sakarya University, Faculty of Computer and Information Sciences, Department of Computer Engineering, evirgen@sakarya.edu.tr

* Corresponding Author: sumeyye@sakarya.edu.tr

Article Info

Received: 4 May, 2018

Accepted: 13 July, 2018

Online: 3 December, 2018

Keywords: artificial neural network; adaptive network fuzzy inference system; student consultancy; course selection

Abstract

Universities offer technical elective courses to allow students to improve themselves in various parts of their majors. Each semester, the students make a decision regarding these technical electives, and the most common expectations students have in this context include, getting education at a better school, getting a better job, and getting higher grades with a view to securing admission into more advanced degree programs. Electing a course on the basis of the interests and skills of the student will naturally translate into achievement. Advisors, in this context, play a major role. Yet, the substantial workload advisors have already assumed prevent them dedicating enough time for exploring the interests and skills of the students, and hence hinder the development of the required relationship between students and their advisors. This study attempts to estimate the achievement level a student intends to elect, on the basis of graduate data received from the database of students of Sakarya University, Faculty of Computer and Information Sciences, and led to the development of a decision-support system. The application used ANFIS and artificial neural network methods among the artificial intelligence techniques, alongside the linear regression model as the mathematical model, whereupon the performance of the methods were compared over the application. In conclusion, it was observed that artificial intelligence techniques provided more relevant results compared to mathematical models, and that, among the artificial intelligence techniques feed forward backpropagation neural network model offered a lower standard deviation compared to ANFIS model.



CrossMark



To cite this article: Kayak, S., Kayak, B. & Evirgen, H. (2018). Student consultancy service: prediction of course grades in course selection phases using artificial intelligence techniques. *Journal of Computer and Education Research*, 6 (12), 142-162. DOI: 10.18009/jcer.421123

Introduction

Universities serve to raise individuals with required social skills, who are capable of solving problems at a national and international scale using their technical skills and qualified to take part in interdisciplinary studies. The curricula are designed in line with this objective. A glance at the curricula of leading Turkish universities (METU, ITU, Boğaziçi etc.) reveals that the first year is dedicated to introductory courses in the field. Such courses are offered as must courses. Students who took the introductory courses are then offered technical electives to allow them improve themselves in their field of interest, and non-technical electives to provide them accumulation of knowledge in different fields and to widen their horizons.

Each semester, students decide on the electives they will take. Their decisions are based on a multitude of factors. However, the factors, or weights thereof, vary from student to student, in line with their expectations from the electives and their differing conceptions of achievement (Babad, 2001). Babad (2001) note the student's characteristics, actual course and instructor characteristics, types of information about courses and situational characteristics (time constraints etc.) among the factors which have an impact on course selection. Kardan, Sadeghi, Ghidary and Sani (2013) on the other hand, mention course characteristics, instructor's characteristics, student's workload, course grade, course type, course time, number of time conflicts, final examination time and student demands among the leading factors concerning course selection (Kardan et al., 2013).

Students may vie for higher grades in order to secure admission to a better school, advance to higher level programs or universities, get a better job, or for personal reasons. Therefore, grading track of the course is a crucial factor affecting the decision of the students in course selection. Indeed, at times learning is relegated to a secondary position for the student, while getting a good grade assumes the primary determining factor position (Kardan et al., 2013). Guessing the potential grade, they would get from the course would provide them guidance in course selection. As noted above, student course selection modelling is a multivariate, nonlinear one (Babad, Ickson & Yelinek, 2008; Babad & Tayeb, 2003; Babad, 2001). Neural network methods, on the other hand, are robust tools for the solution of non-linear problems (Kardan et al., 2013). They enable the construction of models that efficiently describe real world systems, and recently have been successfully applied to approximating nonlinear functions in many disciplines including traditional and online education (Baylari & Montazer, 2009; Kardan et al., 2013; Lo, Chan, & Yeh, 2012;

Giannoukos, Nikolopoulos, Mpardis & Loumos, 2009). A detailed discussion of neural networks is provided in 'Review of existing methods applied in practice'.

Adaptive neural fuzzy inference systems are tools used for the solution of non-linear problems (Babuška & Verbruggen, 2003). They are also used for estimation purposes, in a manner comparable to artificial neural networks (Caner, 2009). A detailed discussion of ANFIS is also provided in 'Review of existing methods applied in practice'.

This study intends to develop a model to provide support for advisor services, on the basis of the estimation of grades students would get from elective courses. Actual data from Sakarya University, Faculty of Computer and Information Sciences, Department of Computer Engineering, gathered in a period of 10 years is used in the study.

The implementation of the application entails four stages. The first one is data collection, and preparation of the dataset containing quantitative and qualitative data, for the application. At this stage, the dataset is subjected to various transformation and filtering processes. The dataset thus prepared contains a wide variety of attributes. That is why the leading attributes in terms of estimation strength need to be identified, on the basis of experts' opinions, and the use of SPSS. Such attributes would then be used as the input parameters of the model. The identification of the technique to provide the best estimation would be the next step of the procedure. The study applies two distinct artificial intelligence techniques, alongside the linear regression model. The aim is to identify the model with the best estimation performance in terms of estimating grades in various technical electives. The final stage, on the other hand, entails the assessment of the estimation performances of the model in various technical electives.

The overall intent is to assist the student, the advisor, the instructor, and even the university administration, by providing a decision support system. Excessive workload on the advisor, and excessive numbers of students assigned to a particular advisor usually lead to a state of affairs whereby the advisors fail to devote sufficient time to individual students. Advisors can check on their screens the courses the student had previously taken; however, they naturally cannot have a complete command of all courses and contents thereof. That is why they may not always reach to an accurate assessment of the courses which would be of benefit to the student, and in which the student would be successful. Thanks to the decision-support system, advisors can have an idea of the compatibility of the contemplated courses with the student, and estimate the achievement level the student will have upon electing the course. The results of the assessment would provide guidance to the advisor. The assessment would also provide insights into the relationships between courses. This would help advisor ascertain any shortcomings of the student. This analysis intends to provide a decision-

support system to help advisor make informed decisions about the student. In a nutshell; such an analyses would help advisor estimate the student's achievement level in the course, and assists her in discovering the abilities of and getting to know the student. The analysis would also provide substantial added value for teachers as well. The instructor aims to move to the upper level to "learning" and "success" according to the needs of the students (Şentürk, 2016). In line with these goals, each semester the instructor will revise or update the syllabus and discuss some of the topics in more detail, and others in less. This analysis may provide insights into the general weaknesses of the class. Estimates can be made about the learning method of the majority of the class. The instructors may have an idea of the class' capabilities in advance, and adjust the course's contents as required. It will facilitate the creation of effective learning environments, one of the primary tasks of teachers (Şahin & Arcagök, 2013). The analysis will also bring substantial benefits to the university administration as well. The administration will be able to project the future capacities of faculties and departments. An observation of the relationships between courses may lead to improvements concerning the number and characteristics of classes. The analysis will also help the university administration with planning for the coming semesters. It would be possible even to get insights into the interrelationships of courses.

Literature Review

The identification of the factors to affect the course election on part of the students, review of their academic performance, estimation of their potential grades in specific courses, estimation of cumulative grade point average at the time of graduation are only a few of the applications of data-mining. The studies aim to enhance achievement levels of the students, provide a better study environment, offer a higher quality education, provide a better educational plan and arrangement for educators and administrators, and do away with the shortcomings of educational processes.

Taylan & Karagözoğlu (2009) used ANFIS method to make estimations of grades students would get in engineering economics courses, on the basis of parameters such as grades of midterms, finals, homework and quizzes, with reference to a dataset of 214 students. Then they compared the results of statistical methods with those of ANFIS method (Taylan & Karagözoğlu, 2009). Oladokun, Adebajo & Charles-Owaba (2008) analyzes the potential performance of students considered for admission. The estimate is based on social, education- and family-related parameters such as age, gender, grades in mathematics, English, chemistry courses, and the location of the university (Oladokun et al., 2008). Zaidah & Daliela (2007) also estimates academic performance of students. His estimates are based on decision trees, artificial neural networks, and linear regression methods. The input

parameters used in his study are the demographic profiles of students, and the cumulative grade point averages the students got in the initial semesters (Zaidah & Daliela, 2007). Şengür & Tekin (2013) used artificial neural networks and decision trees to estimate the CGPA at graduation, for the students of Fırat University, Faculty of Education, Department of Computers and Instructional Technologies Education. The estimates are based on a dataset of 127 students. The input parameters used in that study were the final grades students got through 4 years of undergraduate education (Şengür & Tekin, 2013). Güner & Çomak (2011) made an estimation of the students' level of achievement in mathematics 1 course. The estimate is based on support vector machines. Input parameters used in the study are the data pertaining to students enrolled in 8 distinct departments of Pamukkale University, on the basis of University Entrance Exam scores in 2007. Estimates have a success rate of 86% (Güner & Çomak, 2011). Bozkir, Akcapinar Sezer & Gök (2009) analyzed the factors affecting the level of achievement students have in the University Entrance Exam. Their study is based on the student survey published on OSYM web site with respect to 2008 University Entrance Exam. The analysis is based on decision trees and clustering models (Bozkir et al., 2009). Another analysis by Kardan et al. (2013) reviews the potential factors which have an impact on students' satisfaction levels with the online courses they had elected. The study is based on surveys containing questions on course characteristics, student characteristics, course type, workload of the student, and overlapping schedules. It uses 3-layered backpropagation algorithm. The study attempts to identify factors with an impact on course selection by the students, and the extent of their impact on actual course selection (Kardan et al., 2013). Aher & L. M. R. J. (2012) sought to come up with an optimal combination of the course selection on part of the students, using a combination of clustering algorithm- Simple K-means Algorithm & association rule algorithm- with A Priori Association Rule (Aher & L. M. R. J, 2012). Naser, Zaqout, Ghosh, Atallah & Alajrami (2015) developed a model using artificial neural networks to measure the performance of university second year students. Such factors as high school score, score of subject such as Math I, Math II, Electrical Circuit I, and Electronics I taken during the student freshman year, number of credits passed, student cumulative grade point average of freshman year, types of high school attended and gender were used as input variables for the ANN model. As a result, the ANN model was able to analyze student performance more than 80% correctly. Zacharis (2016) developed a model to predict student performance on a blended learning course environment. The model correctly predicted student performance at 98.3%. Kalejaye, Folorunso, Usman (2015) developed a ANN model for predicting students' academic performance in a University System, based on the previous datasets. The students' CGPAs

from first year through their third year were used as the inputs to train the ANN models. The ANN model predicted the final grade of the students with an accuracy of 91.7%.

Review of Existing Methods Applied in Practice

Estimation of the student's potential academic achievement level in the course she wishes to take is a non-linear problem. Therefore, it is possible to argue forcefully that artificial intelligence techniques in contrast to mathematical models will fare better for such problems. In this study, a mathematical model –linear regression– as well as two machine learning techniques –feed forward backpropagation neural network and adaptive neural fuzzy inference system– were employed in an endeavor to try this argument. On other occasions, these machine learning techniques have been applied successfully to solve various nonlinear problems in two phases: training and testing phases. During the training phase, each technique is fed model input couples, which are then processed through the techniques. Once training is completed, testing stage commences. In the testing stage, data which had not been used in the training stage is fed into the model. Model would then come up with an estimate on the basis of such input. The estimate is the output of the model. The model's estimation performance is an indicator of the technique's applicability. The following sections briefly describe the two machine learning techniques used in this study.

Feed Forward Backpropagation Neural Network

In order to be able to understand the feed forward backpropagation neural network architecture, one should develop a strong grasp of the terms feed forward and backpropagation. The term feed forward signifies that the neural network carries information exclusively in one direction –forward. A simple feed forward neural network is shown in Figure 1. The term “backpropagation” describes how this type of neural network is trained. Backpropagation is essentially a form of supervised training (Heaton, 2008). In this form of learning, network is fed both the input and the corresponding output. The response provided by the network to such inputs are then compared against actual results, with a view to calculating error rate. The weights of the connections among the nodes are then updated with reference to the error rate.

Feed forward backpropagation neural network is basically a technique which forwards information in one direction only –forward–, and seeks improvements in output with reference to the error rate calculation.

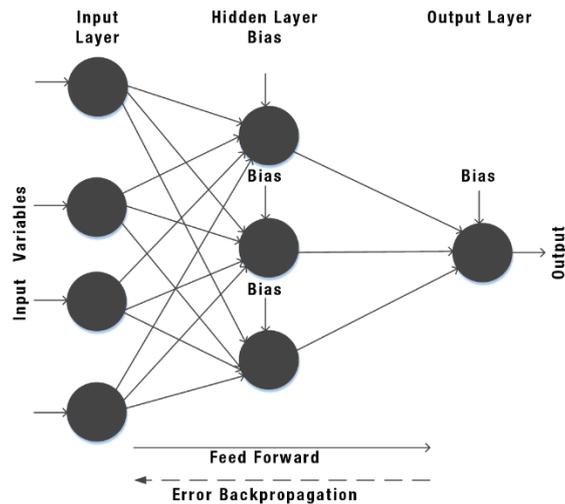


Figure 1. Feed forward backpropagation neural network

Adaptive Neuro-Fuzzy Inference System

ANFIS architecture is developed by embedding the fuzzy inference system into the framework of adaptive networks (Jang, 1993). ANFIS (Adaptive Neuro Fuzzy Inference System) architecture is a hybrid technique based on parallel calculation and learning capabilities of artificial neural networks and the inference capabilities fuzzy logic (Babuška & Verbruggen, 2003). The basic learning rule of ANFIS is the backpropagation gradient descent (Werbos, 1974), which calculates error signals (the derivative of the squared error with respect to each node’s output) recursively from the output layer backward to the input nodes (Jang, 1996). ANFIS architecture is described in Figure 2. Detailed discussion of ANFIS is provided in (Jang, 1993).

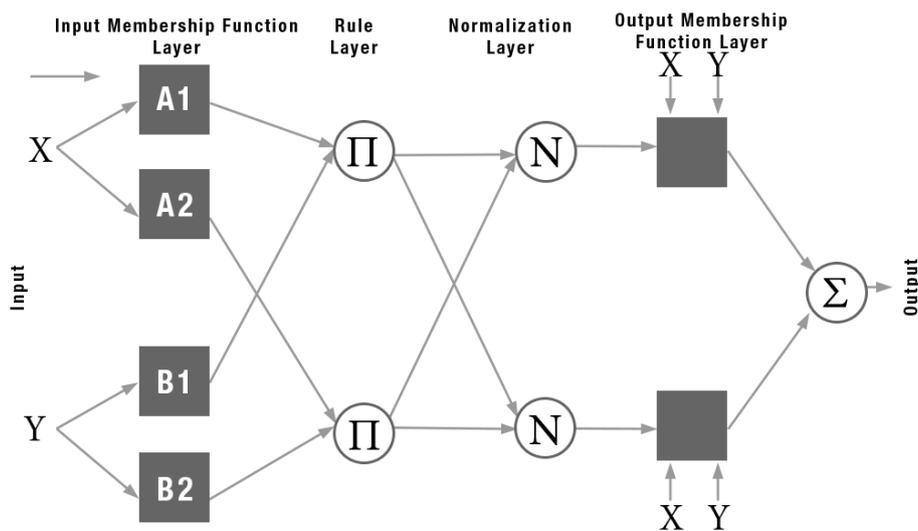


Figure 2. ANFIS architecture

Linear Regression Model

Regression analysis is a method of analysis employed to reveal the relationship

between two or more variables, and providing an insight into the strength of the relationship. Regression analysis can be applied in contexts where one of the variables is the dependent variable, while the rest are independent variables. The dependent variable would be shaped by the independent variables. Independent variables, in turn, have an impact on the dependent one at certain rates, while not getting influenced by the dependent variable in turn. The investigation of the relationship between a single independent variable and a single dependent one is called single-variable regression analysis. In case the number of independent variables is higher than one, the process is called multi-variable regression analysis. The aim of regression analysis is to develop a mathematical model to explain or define the relationship between the dependent variable and one or more independent variables (Seber & Lee Alan J., 2003). The model reveals how dependent variable varies upon changes in conditions (values of the independent variables). This study implements multi-variable regression analysis.

Explanation of Cross Validation Procedure

Cross-validation is a model evaluation method (Najah, El-Shafie, Karim & Jaafar 2010). It allows iterative partitioning of samples into two sets of data; the first set can be used in building the simulation model while the second is used for testing the model (Noureldin et al., 2007). Several methods for implementing the cross-validation theory were proposed in the literature (Noureldin, El-Shafie & Reda Taha, 2007). This study implements a derivative of the cross-validation method: K-fold method. This method divides the whole dataset into two groups: training and testing datasets. Training dataset is used to develop the model, while the testing dataset, which was not fed into the system before, is used to assess the model's performance. This is a universal approach for all cross-validation procedure types. What makes this method stand out, however, is the practice of dividing the set into K subsets. One of the K subsets is used for testing, while the remainder are used for training. Model is trained a total of K times, using different subsets, and is then tested. The average error rate of the model is calculated after each iteration. The mean error rate of the application is the arithmetic average of these error rates. The rate thus calculated serves to signify the model's performance. In this application 192 number of data is divided into 6 subsets in order to have enough data on each sub-set.

Application

Faculty of Computer and Information Sciences of Sakarya University, established in 1994, offers elective courses since year 2000. The elective courses are considered either technical electives, non-technical electives or university-wide common electives.

Technical electives are courses offered specifically for each department. Technical electives build on the foundations laid down in the introductory courses of the department, and seek to provide specialization in various fields. That is why such electives are closely associated with introductory courses. The level of association, however, is not homogenous with all introductory courses. These are the grounds on which this study chose to focus on technical electives.

The applied analysis sought to estimate the final grades the student would get in the technical electives she wished to take. The estimation was made through linear regression analysis and artificial neural networks and ANFIS among leading artificial intelligence techniques. Data from the most popular course of the school –Data Mining– was used to identify the method with the best estimation performance. The model thus developed was then applied to estimate the achievement levels in technical electives selected in random: BSM421 (Computer Graphics), BSM435 (Internet Engineering) and BSM425 (Artificial Intelligence). The model's performance in those courses was analyzed. The model's performance in those courses was analyzed using MSSQL 2012, Excel 2013 and SPSS. For the actual implementation of the techniques, Matlab 2012 platform was employed.

Data Gathering and Definition of the Data Set

The data pertaining to the technical electives used for the application were obtained from the Student Information Database of Sakarya University. Then, data was processed through a C# application, to make it importable to Excel.

In order to render the dataset usable, certain editing steps were applied. Gender and achievement parameters were rendered numerical. Another piece of editing was filtering the dataset for the data which met all the requirements. The gender parameter was shown with parameters 1 and 2 denoting female and male students respectively. The achievement parameter, on the other hand, was defined in a range of 1 to 8 (AA, BA, BB, CB, CC, DC, DD, FF). At Sakarya University, highest grade is AA (1), while the lowest passing grade is DD (7). FF (8), on the other hand, indicates that the student failed the course. The grade point average for the semester is shown in the range of 0-4. No conversion was required for this parameter. Filtering of the data which meet all the requirements was achieved by removing the data which do not meet all the parameters, from the dataset. This is effected with a view to preventing missing values leading to a deviation in the final result. This process led to a significant reduction in the dataset.

Identifying Input Variables

The data cluster contains gender, year of birth, year of admission, and transcript

details of the student. Such a wide range of data items is simply too much for the model to handle in a useful manner. Therefore, one should filter for the pieces of data with the highest possible impact on the result.

Sakarya University Education Information System contains the contents, pre-requisites and instructional details of all the courses. The identification of the input parameters was based on a review of the details in the information system, and consultations with the instructors. The data thus gathered were reviewed to analyse the courses believed to have an impact on the course subject to the estimation (BSM 429, BSM 421, BSM425 and BSM 435), as well as the gender and grade point average figures were examined through correlation analysis in SPSS.

The Table 1, presents the impact of 8 variables on the performance for BSM429. As seen in Table 1, OD6 (grade point average of a student at the end of 6th semester) and BSM210 (Discrete Computational Structures) are the two features affecting the result most. While OD6 had a negative impact, given its nature, the impact of the BSM210 variable was positive. MAT217 (Numerical Analysis) is the third most significant variable, following closely behind. These variables were used in the construction of input parameters.

Table 1. BSM429 correlation analysis

		<i>BSM429</i>	<i>Sex</i>	<i>OD6</i>	<i>BSM103</i>	<i>BSM101</i>	<i>BSM102</i>	<i>BSM210</i>	<i>MAT217</i>	<i>IST108</i>
Pearson Correlation	<i>BSM429</i>	1,000	-,293	-,577	,089	,259	,079	,528	,399	,249
	<i>Sex</i>	-	1,000	,061	,169	,033	,155	-,286	,027	,006
	<i>OD6</i>	-	-	1,000	-,315	-,434	-,164	-,507	-,384	-,380
	<i>BSM103</i>	-	-	-	1,000	,111	,256	,013	,000	,112
	<i>BSM101</i>	-	-	-	-	1,000	-,024	,168	,319	,266
	<i>BSM102</i>	-	-	-	-	-	1,000	,034	-,019	,024
	<i>BSM210</i>	-	-	-	-	-	-	1,000	,399	,141
	<i>MAT217</i>	-	-	-	-	-	-	-	1,000	,137
	<i>IST108</i>	-	-	-	-	-	-	-	-	1,000
Sig.(1- tailed)	<i>BSM429</i>	.	,000	,000	,110	,000	,138	,000	,000	,000
	<i>Sex</i>	-	.	,201	,009	,326	,016	,000	,355	,466
	<i>OD6</i>	-	-	.	,000	,000	,012	,000	,000	,000
	<i>BSM103</i>	-	-	-	.	,063	,000	,428	,499	,062
	<i>BSM101</i>	-	-	-	-	.	,370	,010	,000	,000
	<i>BSM102</i>	-	-	-	-	-	.	,321	,397	,371
	<i>BSM210</i>	-	-	-	-	-	-	.	,000	,026
	<i>MAT217</i>	-	-	-	-	-	-	-	.	,029
	<i>IST108</i>	-	-	-	-	-	-	-	-	.

The review of the data pertaining to BSM421 (Computer Graphics) in the educational information system revealed, with reference to the views of the instructors and students as well, that BSM421 was a course associated with BSM207 (Data Structures), BSM102 (Algorithm and Programming 2), BSM103 (Algorithm and Programming 3), Mat113 (Linear Algebra), FIZ111 (Physics 1), FIZ112 (Physics 2), MAT111 (Mathematics 1), MAT112

(Mathematics 2) and BSM210 (Discrete Operational Structures). Correlation analysis was applied to support this view. Gender and OD6 parameters were also included in the analysis. A glance at Table 2 reveals that OD6 (grade point average of a student as of the end of the 6th semester), BSM207 (Data Structures) and MAT113 (Linear Algebra) had the largest impact on the result. That is why these characteristics were used as input parameters for the estimation for the course BSM421.

Table 2. BSM421 correlation analysis

		<i>BSM421</i>	<i>Sex</i>	<i>OD6</i>	<i>BSM102</i>	<i>BSM103</i>	<i>BSM207</i>	<i>BSM210</i>	<i>MAT113</i>
Pearson Correlation	<i>BSM421</i>	1,000	-,053	-,590	,190	,175	,371	,193	,315
	<i>Sex</i>	-	1,000	,017	,114	,101	-,163	-,161	,097
	<i>OD6</i>	-	-	1,000	-,268	-,267	-,542	-,461	-,392
	<i>BSM102</i>	-	-	-	1,000	,382	,170	,145	,055
	<i>BSM103</i>	-	-	-	-	1,000	,184	,048	,086
	<i>BSM207</i>	-	-	-	-	-	1,000	,261	,241
	<i>BSM210</i>	-	-	-	-	-	-	1,000	,298
	<i>MAT113</i>	-	-	-	-	-	-	-	1,000
	Sig. (1-tailed)	<i>BSM421</i>	.	,238	,000	,005	,009	,000	,004
<i>Sex</i>		-	.	,409	,063	,087	,014	,015	,096
<i>OD6</i>		-	-	.	,000	,000	,000	,000	,000
<i>BSM102</i>		-	-	-	.	,000	,011	,025	,228
<i>BSM103</i>		-	-	-	-	.	,006	,261	,125
<i>BSM207</i>		-	-	-	-	-	.	,000	,001
<i>BSM210</i>		-	-	-	-	-	-	.	,000
<i>MAT113</i>		-	-	-	-	-	-	-	.

The course contents and pre-requisites concerning BSM425 were analyzed over the educational information systems. In consultations with different instructors and students, course parameters to play an active role in the estimation for the course were found to be BSM103 (Discrete Operational Structures), BSM207 (Data Structures), BSM306 (Software Engineering) and MAT217 (Digital Analysis). To support this finding, Gender and OD6 parameters were also included in the correlation analysis. The analysis result thus obtained reveals that OD6 has a high correlation factor. It is followed by BSM207 and BSM306. The results of the correlation analysis are provided in Table 3. The final grade estimation for the course BSM425 will be based on OD6, BSM207 and BSM306 parameters.

Table 3. BSM425 correlation analysis

		BSM425	BSM103	BSM207	BSM306	MAT217	Sex	OD6
Pearson Correlation	BSM425	1,000	,162	,215	,284	,056	-,008	-,475
	BSM103	-	1,000	,117	,269	,178	,009	-,509
	BSM207	-	-	1,000	,036	,134	-,138	-,368
	BSM306	-	-	-	1,000	,306	-,003	-,404
	MAT217	-	-	-	-	1,000	,086	-,492
	Sex	-	-	-	-	-	1,000	,052
	OD6	-	-	-	-	-	-	1,000
Sig. (1-tailed)	BSM425	.	,074	,027	,005	,310	,471	,000
	BSM103	-	.	,150	,008	,056	,469	,000
	BSM207	-	-	.	,374	,117	,110	,000
	BSM306	-	-	-	.	,003	,488	,000
	MAT217	-	-	-	-	.	,223	,000
	Sex	-	-	-	-	-	.	,000
	OD6	-	-	-	-	-	-	.

BSM435 (Internet Engineering) course was reviewed over the educational information system. The prerequisites for the course are Data Communications and Computer Networks. That is why, these courses are deemed important to note among input parameters. Details of other courses were also reviewed. However, no potentially compatible course was identified. Correlation analysis sought to investigate the impact of the data communications, computer networks, gender and OD6 parameters on the course. The results of the analysis are presented in Table 4. A glance at Table 4 reveals that BSM304, BSM305 and OD6 are the parameters which exhibit the highest correlation factor. That is why these were registered as input parameters.

Table 4. BSM435 correlation analysis

		BSM435	Sex	BSM305	BSM304	OD6
Pearson Correlation	BSM435	1,000	,108	,394	,397	-,660
	Sex	-	1,000	,016	-,117	,030
	BSM305	-	-	1,000	,309	-,493
	BSM304	-	-	-	1,000	-,490
	OD6	-	-	-	-	1,000
Sig. (1-tailed)	BSM435	.	,100	,000	,000	,000
	Sex	-	.	,426	,084	,364
	BSM305	-	-	.	,000	,000
	BSM304	-	-	-	.	,000
	OD6	-	-	-	-	.

Estimation of Students' Level of Academic Success in the "Data Mining" Course Through the Feed Forward Back Propagation Neural Network Method

Feed forward backpropagation neural network method is capable of learning and generalizations. This makes it capable of successful estimations. The dataset forms the input parameters of this technique. Feed forward backpropagation neural network offers a

multitude of settings to fine-tune its characteristics. In this study, various settings were tried and the network offering the best performance was implemented. The network is essentially a two-layered structure. The functions used to train the network are TRAINLM, LEARNLDM, MSE and TANSIG for Training, Adaption learning, Performance and Transfer functions respectively. The analysis used 160 data items to train the network, and 32 previously unused data items to make performance assessment. The network's performance assessment is effected through cross-validation method.

First of all, the whole dataset was divided into 6 subsets. 5 out of the 6 subsets were used for training, and 1 was used for the testing of the network. 6 distinct settings were applied to reach all possible training and test combinations. In each setting, the network was trained using distinct subsets, and tested with an unused subset. Then training performances for all tested states are calculated. Table 5, shows the folds used in the training and the testing of the model, as well as training performances, deviation of estimations from actual figures, and the frequencies of such deviations, one potential setting at a time. In sum, 123 out of a total of 192 data items were estimated accurately, while 64 were estimated with a deviation of 1 unit, and 5 were estimated with a deviation of 2 units. Each unit of deviation represents one step of change in the letter grade, for better or worse. For instance, if the student received a grade of BB in a given course, 1 unit of deviation in the estimation by the system would produce a grade of BA or CB.

Table 5. K Fold Cross Validation Error Results for Feed Forward BackPropagation Neural Network

<i>Fold Used in Training Phases</i>	<i>Folds Used in Testing Phases</i>	<i>Deviation:1 unit</i>	<i>Deviation: 2 unit</i>	<i>Correct number</i>
1,2,3,4,5	6	14	0	18
1,2,3,4,6	5	13	0	19
1,2,3,5,6	4	7	1	24
1,2,4,5,6	3	11	1	20
1,3,4,5,6	2	12	1	19
2,3,4,5,6	1	7	2	23
1,2,3,4,5	6	14	0	18

Prediction of Academic Success in Data Mining Course using ANFIS

Adaptive neural fuzzy inference system is a multilayer feed forward network which uses neural network learning algorithms and fuzzy reasoning to map an input space to an output space. ANFIS possesses a substantial capacity of learning, constructing, expensing and classifying (Yan, Zou & Wang, 2010). These characteristics make ANFIS model a strong tool in successful estimation applications. The set up of the ANFIS model, just like the feed

forward backpropagation neural network, offers various settings. A multitude of different settings were tried, and the best performing one was used in this analysis. When developing the ANFIS model, first of all membership functions were assigned for each input. Within the framework of the ANFIS model, 1 membership function was assigned to the gender input, 2 membership function was assigned to the cumulative grade point average by the end of the 6th semester, and 3 membership function was assigned to BSM201 and MAT217 courses. Gbellmf was used for inputs, while linear membership function was used for outputs. Hybrid learning rule was chosen as the learning rule. The ANFIS model thus developed was trained using the dataset, and subjected to performance assessment using the dataset which was not applied before. 160 out of a total of 192 data items were used to train the model, and 32 were used for performance assessment. The model's performance assessment is effected through cross-validation procedure. The whole dataset was divided into 6 subsets, with 1 subset containing 32 data items was set aside for performance assessment. The Table 6 shows the folds used in the training and the testing of the model, as well as deviation of estimations from actual figures, and the frequencies of such deviations, one potential setting at a time.

Table 6. K Fold Cross Validation Error Results for ANFIS

<i>Fold Used in Training Phases</i>	<i>Folds Used in Testing Phases</i>	<i>Deviation:1 unit</i>	<i>Deviation: 2 unit</i>	<i>Correct number</i>
1,2,3,4,5	6	14	1	17
1,2,3,4,6	5	10	1	21
1,2,3,5,6	4	12	1	19
1,2,4,5,6	3	12	3	17
1,3,4,5,6	2	11	1	20
2,3,4,5,6	1	12	1	19
1,2,3,4,5	6	14	1	17

Prediction of Academic Success in Data Mining Course using Multiple Linear Regression Model

In order to describe the relationship between the input parameters and the output of the "Data Mining" course, the linear regression mathematical model was tried. Linear regression analysis represents the relationship between the dependent variable and the independent variable(s) through a linear function. The coefficients table produced through the regression analysis is provided below.

The Table 7 shows the regression factors used for the regression equilibrium, and the significance levels thereof. B column presents the coefficients for the multiple regression formula. Beta column, on the other hand, shows to what extent the model would be affected

in response to a one-unit increase in independent variables. Sig. column shows the statistical significance of each independent variable in terms of its relationship with the dependent one. As Table 7 reveals, the multiple regression analysis is based on the following statement:

$$-0.900*Gender-1.382*OD6-0.193*BSM210+0.191*MAT217+6.514$$

Table 7. Coefficients

	Unstandardized Coefficients		Standardized Coefficients		95,0% Confidence Interval for B			Correlations		
	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
<i>(Constant)</i>	6,514	,800		8,141	,000	4,933	8,094			
<i>Gender</i>	-,900	,242	-,230	-3,718	,000	-1,378	-,422	-,288	-,286	-,216
<i>GPA6</i>	-1,382	,233	-,406	-5,943	,000	-1,841	-,923	-,574	-,431	-,346
<i>BSM210</i>	,193	,079	,178	2,435	,016	,036	,349	,519	,192	,142
<i>MAT217</i>	,191	,063	,201	3,051	,003	,067	,315	,416	,238	,177

When the formula was used with the test data, accuracy of the results was 0.1875. This is a fairly low figure compared to the results of the artificial intelligence techniques employed. That is why the study focused more on ANN and ANFIS artificial intelligence models.

Model Selection

This section discusses the applicability of ANFIS method and feed forward backpropagation neural network technique for the estimation of the student's academic success in courses she intends to take, and compares the performance of the two techniques in applied scenarios. The application case was the Data Mining course, a technical elective offered at Sakarya University, Faculty of Computer and Information Sciences. A total of 192 data items were used in the analysis. ANFIS model estimated 113 out of a total of 192 data items (59\%) accurately, while 71 were estimated with a deviation of 1 unit, and 8 were estimated with a deviation of 2 units. Feed forward backpropagation neural network technique, on the other hand estimated 123 out of a total of 192 data items (64\%) accurately, while 64 were estimated with a deviation of 1 unit, and 5 were estimated with a deviation of 2 units. The standard deviation of ANFIS model was 0.729, while that of the feed forward backpropagation neural network model was 0.660.

Accurate estimation of the grade each student would get is no simple task, given the existence of parameters which cannot be represented in numbers or assessed properly, but which, nonetheless affect the estimation. That is why providing the responses for individual students in intervals rather than exact grades would be a wiser choice. The feed forward

backpropagation neural network technique, in this context, is considered a better option to be applied in comparison to ANFIS model.

The Application of ANN Method on Different Technical Electives

Data Mining course was chosen as the object of comparison of the achievement estimation performance of feed forward artificial neural networks and ANFIS method. It was observed that feed forward artificial neural networks performed better than the ANFIS method. On the basis of this finding, a feed forward artificial neural network model was developed for each and every one of BSM421, BSM425 and BSM435 courses, and the model was run to make estimations.

As noted above in analysis was run for each and every one of BSM421, BSM425 and BSM435 courses, and input parameters were established. The datasets comprising input parameters were subjected to various conversion and filtering procedures described in “Data gathering and definition of the data set” section. In conclusion of all such analyses and procedures, BSM421, BSM425, and BSM435 courses were left with datasets of 196, 149 and 137 items respectively.

Cross-validity method is applied. The datasets for BSM421, BSM425 and BSM435 courses are divided into 4 subsets (folds). 1 fold is reserved for testing, while the remaining 3 are used for training.

Table 8 shows the folds used in the development of the model, the fold used for the testing of the model, and the deviation of the estimates from actual figures, alongside deviation frequencies, with each line representing a potential case for the course BSM421.

Table 8. The performance displayed by the ANN model for the BSM421 course

Fold Used in Training Phases	Folds Used in Testing Phases	Deviation:1 unit	Deviation: 2 unit	Correct number
2,3,4	1	17	0	32
1,3,4	2	17	2	30
1,2,4	3	23	0	26
1,2,3	4	13	1	35

As Table 8, suggests the feed forward backpropagation network estimated 123 out of 196 data items (63%) accurately. 70 data items were estimated with a deviation of 1 just unit, and 3 were estimated with 2.

Table 9 shows the folds used in the development of the model, the fold used for the testing of the model, and the deviation of the estimates from actual figures, alongside deviation frequencies, with each line representing a potential case for the course BSM425.

As Table 9, suggests ANN model estimated 96 out of 149 data items (64%) accurately. 44 data items were estimated with a deviation of 1 just unit, and 9 were estimated with 2.

Table 9. The performance displayed by the ANN model for the BSM425 course

Fold Used in Training Phases	Folds Used in Testing Phases	Deviation:1 unit	Deviation: 2 unit	Correct number
2,3,4	1	9	2	26
1,3,4	2	12	2	23
1,2,4	3	14	3	21
1,2,3	4	9	2	26

Table 10 shows the folds used in the development of the model, the fold used for the testing of the model, and the deviation of the estimates from actual figures, alongside deviation frequencies, with each line representing a potential case for the course BSM435.

Table 10. The performance displayed by the ANN model for the BSM435 course

Fold Used in Training Phases	Folds Used in Testing Phases	Deviation:1 unit	Deviation: 2 unit	Correct number
2,3,4	1	11	0	23
1,3,4	2	10	0	24
1,2,4	3	12	0	22
1,2,3	4	13	1	21

As Table 10, suggests, ANN model estimated 90 out of 137 data items (66%) accurately. 46 data items were estimated with a deviation of 1 just unit, and 1 were estimated with 2.

Conclusion and Suggestions

Feed forward artificial neural network was used to estimate the grades the students would get in 4 technical elective courses. This model developed for estimation led to acceptable results for the courses in question. The model's estimates did not deviate from the actual figures by more than 2 levels. The review of estimates stipulating deviation revealed that the results of the artificial neural network model provided a more or less accurate and significant picture of the potential success the students will have in the course. The applied analysis sought to estimate the final grades the student would get in the technical electives she wished to take. In contrast to previous studies, input parameters were identified using SPSS, through the application of pearson-correlation. First of all, the relationships between the courses were analysed. Then a comparison between a mathematical model –linear regression analysis– was compared against two artificial intelligence models –ANN and ANFIS. The present study intends to develop a decision-support system which can be applied at a university level over a web application. However, significant efforts are required for the development of the model. Thereafter, the model would produce the results automatically, provided that required level of integration is established. It is possible to

estimate the student's performance online, with reference to the intended course registration. The steps required for the implementation of the system at an actual university can be summarized as follows:

- The technical electives to be included in the framework of this analysis should be selected. The most important factor in this context is the availability of sufficient data for the courses in question.
- For each such course, the course coordinators would then be required to specify courses with potential associations. Throughout the whole process, this is the only step which requires user interaction. The system then applies a correlation analysis to identify the most significant ones among the courses, and labels them input parameters.
- The identified input parameters would then be used to create the model for each course automatically. Once integrated with the advisor-information system, the model would present its output to the advisors, for each student.

The provision of such a service at the university level is among the most important contributions such a model may have. We are confident that such a service would shed light for the scientists working on such matters.

The system can be used as a tool to assist advisory and counselling services, on the basis of past data. This service can be used to direct students to courses with which they may have better success. Of course, students cannot always be expected to choose the course will have better success. A result not be satisfied with estimates; it does not mean that students will not definitely choose that course. The result also means that student should be put more effort and remedy their deficiencies. In this study, we aimed to shed light on the students benefiting from past experiences. Secret correlations can be seen in courses with the analysis. The advisors, on the other hand, may find an assistant for course recommendations in the system. In the context of using a trained network as such a tool, the implementation of a web-based application available for students and advisors would be valuable. In particular, this would provide a unique input in the course registration processes. A screenshot is offered as an example in Figure 3.

In practice, a specific model will be required to be developed for each course. The development of the model, on the other hand, will require historical data for the course. For obvious reasons, the system cannot be used for a brand new course. The model for a course

with sufficient past data is developed through an offline learning process. The next step of the endeavor may lie in the implementation of an online-learning process for the development of models.

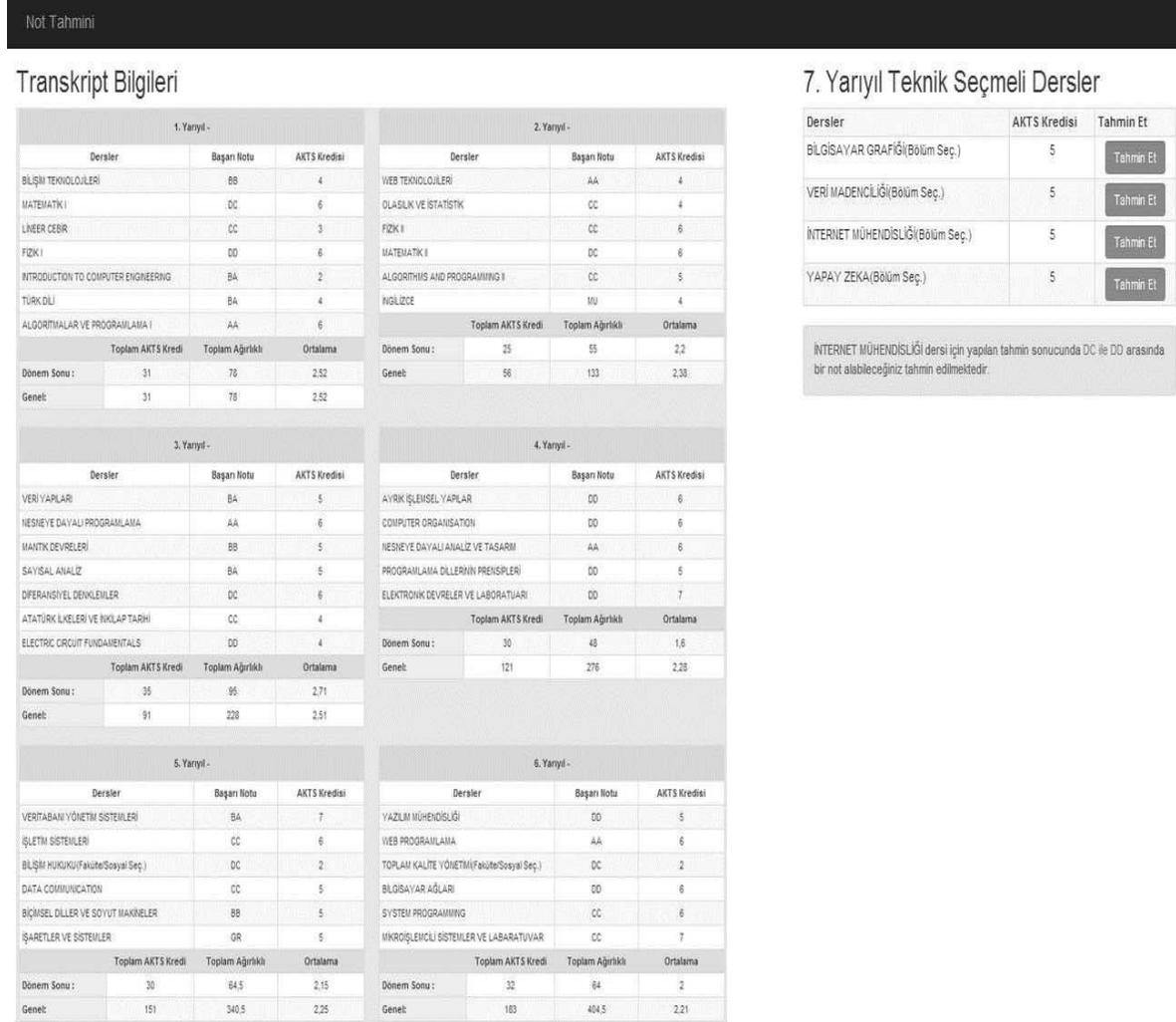


Figure 3. Example application screenshot

References

- Aher, S., & L. M. R. J, L. (2012). A comparative study of association rule algorithms for course recommender system in e-learning. *International Journal of Computer Applications*, 48-52
- Babad, E. (2001). Students' course selection: differential considerations for first and last course students' course selection: Differential Considerations for First and Last Course, 42(4), 469–492. Retrieved from <http://www.jstor.org/stable/30069473>
- Babad, E., Ickson, T., & Yelinek, Y. (2008). Antecedents and correlates of course cancellation in a university "drop and add" period. *Research in Higher Education*, 49(4), 293–316. <http://doi.org/10.1007/s11162-007-9082-3>
- Babad, E., & Tayeb, A. (2003). Experimental analysis of students' course selection. *The British Journal of Educational Psychology*, 73(Pt 3), 373–393. <http://doi.org/Doi.10.1348/000709903322275894>

- Babuška, R., & Verbruggen, H. (2003). Neuro-fuzzy methods for nonlinear system identification. *Annual Reviews in Control*, 27(1), 73–85. [http://doi.org/10.1016/S1367-5788\(03\)00009-9](http://doi.org/10.1016/S1367-5788(03)00009-9)
- Baylari, A., & Montazer, G. a. (2009). Design a personalized e-learning system based on item response theory and artificial neural network approach. *Expert Systems with Applications*, 36(4), 8013–8021. <http://doi.org/10.1016/j.eswa.2008.10.080>
- Bozkir, A., Akcapinar Sezer, E., & Gök, B. (2009). Öğrenci seçme sınavında (öss) öğrenci başarımını etkileyen faktörlerin veri madenciliği yöntemleriyle tespiti. Uluslararası İleri Teknolojiler Sempozyumu (IATS'09)
- Caner, M. (2009). Estimation of specific energy factor in marble cutting process using ANFIS and ANN, 221–226.
- Güner, N., & Çomak, E. (2011). Mühendislik öğrencilerinin matematik i derslerindeki başarısının destek vektör makineleri kullanılarak tahmin edilmesi. *Pamukkale Univ Muh Bilim Dergisi*, 87-96
- Heaton, J. (2008). *Introduction to neural networks for C# (2 edition)*. Heaton Research, Incorporated.
- Jang, J. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665–685. <http://doi.org/10.1109/21.256541>
- Jang, J. (1996). *Input selection for ANFIS learning*. *Fuzzy Systems*, Proceedings of the Fifth. 1493–1499. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=552396
- Kalejaye, B., Folorunso, O., & Usman, O. (2015). Predicting students' grade scores using training functions of artificial neural. *Journal of Natural Science, Engineering and Technology*
- Kardan, A. a., Sadeghi, H., Ghidary, S. S., & Sani, M. R. F. (2013). Prediction of student course selection in online higher education institutes using neural network. *Computers & Education*, 65, 1–11. <http://doi.org/10.1016/j.compedu.2013.01.015>
- Lo, J. J., Chan, Y. C., & Yeh, S. W. (2012). Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Computers and Education*, 58(1), 209–222. <http://doi.org/10.1016/j.compedu.2011.08.018>
- Lykourantzou, I., Giannoukos, I., Mpardis, G., Nikolopoulos, V., & Loumos, V. (2009). Early and dynamic student achievement prediction in E-learning courses using neural networks. *Journal of the American Society for Information Science and Technology*, 60(2), 372–380. <http://doi.org/10.1002/asi.20970>
- Lykourantzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers and Education*, 53(3), 950–965. <http://doi.org/10.1016/j.compedu.2009.05.010>
- Najah, A. a., El-Shafie, A., Karim, O. a., & Jaafar, O. (2010). Water quality prediction model utilizing integrated wavelet-ANFIS model with cross-validation. *Neural Computing and Applications*, 21(5), 833–841. <http://doi.org/10.1007/s00521-010-0486-1>

- Naser, S., Zaqout, I., Ghosh, M., Atallah, R., & Alajrami, E. (2015). Predicting student performance using artificial neural network: in the faculty of engineering and information technology. *International Journal of Hybrid Information Technology*, 221-228
- Noureldin, A., El-Shafie, A., & Reda Taha, M. (2007). Optimizing neuro-fuzzy modules for data fusion of vehicular navigation systems using temporal cross-validation. *Engineering Applications of Artificial Intelligence*, 20(1), 49-61. <http://doi.org/10.1016/j.engappai.2006.03.002>
- Oladokun, V., Adebajo, A., Sc, B., & Charles-Owaba, O. (2008). Predicting Students ' academic performance using artificial neural network : a case study of an engineering course. *The Pacific Journal of Science and Technology*, 72-79
- Seber, G. A. F., & Lee Alan J. (2003). *Linear regression analysis*. Wiley-Interscience.
- Şahin, Ç., & Arcagök, S. (2013). İlköğretim öğretmenlerinin eğitim araştırmalarına yönelik yaklaşımları. *Journal of Computer and Education Research*, 1(2), 1-20.
- Şengür, D., & Tekin, A. (2013). Prediction of student's grade point average by using the data mining methods. *Bilişim Teknolojileri Dergisi*, 6(3), 7-16.
- Şentürk, M. (2016). Sosyal bilgiler dersinde işbirlikli öğrenme yöntemlerinin akademik başarı üzerindeki etkisi. *Journal of Computer and Education Research*, 4(8), 205-221.
- Taylan, O., & Karagözoğlu, B. (2009). An adaptive neuro-fuzzy model for prediction of student's academic performance. *Computers & Industrial Engineering*, , 732-741.
- Werbos, P. J. (1974). Beyond regression: new tools for prediction and analysis in the behavioral sciences foundations. Retrieved from <http://www.citeulike.org/group/1938/article/1055600>
- Yan, H., Zou, Z., & Wang, H. (2010). Adaptive neuro fuzzy inference system for classification of water quality status. *Journal of Environmental Sciences*, 22(12), 1891-1896. [http://doi.org/10.1016/S1001-0742\(09\)60335-1](http://doi.org/10.1016/S1001-0742(09)60335-1)
- Zacharis, N. (2016). Predicting student academic performance in blended learning using artificial neural. *International Journal of Artificial Intelligence and Applications (IJAlA)*
- Zaidah, I., & Daliela, R. (2007). Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression. 21st Annual SAS Malaysia Forum.