DOCUMENT RESUME

ED 358 169 TM 019 973

AUTHOR Song, Qiang; Chissom, Brad S.
TITLE New Models for Forecasting Enrollments: Fuzzy Time Series and Neural Network Approaches.
PUB DATE Apr 93
NOTE 27p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Enrollment Projections; Enrollment Trends; *Evaluation Methods; Higher Education; *Mathematical Models; Predictive Measurement; *Research Methodology; Robustness (Statistics); School Statistics; Trend Analysis; Universities
IDENTIFIERS *Artificial Neural Networks; *Fuzzy Time Series; Stochastic Analysis

ABSTRACT

Since university enrollment forecasting is very important, many different methods and models have been proposed by researchers. Two new methods for enrollment forecasting are introduced: (1) the fuzzy time series model; and (2) the artificial neural networks model. Fuzzy time series has been proposed to deal with forecasting problems within a fuzzy environment. In this model, the uncertainty encountered in the forecasting process is taken as being produced by our incomplete understanding of nature. As a result, it is different from any stochastic methods. The major problem with this method is that the forecasted values depend to some degree on our interpretations of the outputs of the forecasting model, which makes the process quite subjective. Artificial neural networks represent an advanced technology applied in engineering. When applied in forecasting, uncertainty is ignored. Consequently, the model is a deterministic one. In spite of this, because of the ability to generalize once trained, the network model has robustness. Examples show how these methods are applied. Six figures and two tables illustrate these applications. (SLD)

New Models for Forecasting Enrollments:

Fuzzy Time Series and Neural Network Approaches

Qiang Song and Brad S. Chissom

Area of Behavioral Studies

College of Education

The University of Alabama

Paper presented at the meeting of the American Educational Research Association;

Atlanta, GA; April 12-16, 1993

Abstract

Since university enrollment forecasting is very important, many different methods and models have been proposed by researchers. This paper introduces two new models for enrollment forecasting: the fuzzy time series model and the artificial neural networks model. Fuzzy time series has been proposed to deal with forecasting problems under a fuzzy environment. In this model, the uncertainty encountered in the forecasting process is taken as being produced by our incomplete understanding of nature. Hence, it is different from any stochastic methods. Artificial neural networks are new technology applied in engineering. When applied in forecasting, uncertainty is ignored and thus the model is a deterministic one. In spite of this, because of the ability to generalize once trained, the network model has robustness. Through examples, this paper will indicate how these models are applied.

1. Introduction

It is of importance to have reasonably accurate estimate of the future enrollment for a university. For this reason, researchers have proposed many different methods for forecasting enrollments. Many factors influence university enrollments as can be seen from the different models developed by various authors. The key point of developing a forecasting model is to clarify the important factors and their relationships with enrollments. Different factors and relationships will lead to different models or methods. Classifications of, and comments on forecasting methods and models can be found in Gardner (1981) and Shaw (1984).

Hoenack and Weiler (1979) developed a forecasting model which had 11 equations and considered 19 factors. The model was used for a case application with good accuracy. This model was, in essence, an econometrics model used primarily in economic analysis.

Weiler (1980) used growth curve models, often used in the analysis of the sales of newly developed products, to forecast enrollments for the short-period between the beginning of the application and the start of the fall semester. The models were used to forecast enrollments of the University of Minnesota with forecasting errors* ranging from 2.4% to 16%, with an average error** of 9.7%.

The double exponential smoothing method can be used to forecast enrollments by applying a weighing factor. Obviously, the main weakness of this method is the question of how to determine an appropriate value for the weight. Gardner (1981) discussed how to select this weight factor.

Warrack and Russell (1983) developed a forecasting model using a motivational index and a demand index. These indices were tested with surveys of more than 13,000 high school students and validated with respect to actual outcomes by surveying more than 5,000 students about their post-high school plans. Forecasted enrollments of the four universities considered in their study had forecasting errors that ranged from 3.7% to 14.4%. The main drawback of this model is the costly survey which is its foundation.

Pope and Evans (1985) developed a Decision Support System for forecasting enrollments. In their study, the enrolling freshmen were classified into four divisions, and for each division, a different model was derived. However, forecasting errors were not reported, precluding comparisons with errors from other models.

Chatman (1986) applied regression to the accumulative number of accepted students to forecast enrollments. Not only the number, but also the scholastic ability (SAT scores) of the enrolling students could be forecasted. His study yielded forecasting errors ranging from 2.8% to 23%.

Paulsen (1989) proposed a step-by-step method that was conceptually similar to those by Weiler (1980), Pope and Evans (1985), and Chatman (1986). This method achieved an average forecasting error of 2.4% which is the smallest of all the short-term forecasting models reviewed in this paper.

From the literature review, one can perceive that researchers have been looking for new methods to achieve a unique goal: forecasting enrollment for universities. Therefore, the purpose of this paper is to introduce two new models for enrollment forecasting: the fuzzy time series model and the artificial neural networks model.

The fuzzy time series model was developed based on fuzzy logic invented by Dr. Zadeh (Zadeh, 1965,1975a-d) of The University of California at Berkeley. As we know it, any forecasting model utilizes historical data to make a good guess at future events. Because of the noise or uncertainty existing in the data, any prediction of the future also has this kind of uncertainty in it. In stochastic models, the uncertainty is treated as being caused by random factors, while in the fuzzy time series model the uncertainty is taken as being caused by our incomplete understanding of nature. Thus, historical data can

* Forecasting error=(|forecasted value - actual value|/actual value)*100%

** Average forecasting error=(sum of forecasting errors)/(total # of errors) be fuzzy. Since historical data have fuzzy characteristics, nothing but fuzzy logic can be applied. Dealing with fuzziness of historical data is the most important characteristics of fuzzy time series models.

Artificial neural networks are an advanced technology applied in engineering and have achieved fruitful application results (Freeman et al, 1991). Yet, in educational research, we have not found any applications of this technology, and it is our belief that neural networks will be a competitive model in forecasting enrollments if we assume that neural networks can function as a mapping between influencing factors and enrollments. To verify this, we will use a 3-layer back-propagation network as the forecasting model using original data, and explore the possibility of improving the results by deriving a simple equation for forecasting.

2. Fuzzy Time Series Models

2.1 Fuzzy time series

Fuzzy time series was proposed by the authors (Song and Chissom, 1993a,1993b) to deal with forecasting problems in which historical data are fuzzy sets or linguistic values. Generally, a fuzzy time series can be defined as follows:

Fuzzy Time Series. Let $Y(t)$ ($t=.....,0,1,2,......$), a subset of R^1, be the universe of discourse on which fuzzy sets $f_i(t)$ ($i=1,2, ...$) are defined and $F(t)$ is a collection of $f_1(t)$, $f_2(t)$,... ..., then $F(t)$ is called a fuzzy time series defined on $Y(t)$ ($t=.....,0,1,2,....$).

In the above definition, $F(t)$ is a linguistic variable and $f_i(t)$ ($i=1,2,.....$) is the possible linguistic values of $F(t)$. It can be seen that no hints can be found explicitly from the definition about how to model fuzzy time series, thus leaving the possibility of different models. Of all fuzzy time series models, the simplest one is the first-order time-invariant model, which can be expressed as:

$$F(t)=F(t-1)\circ R(t,t-1)=F(t-1)oR(t,t-1) \qquad (1)$$

where $R(t,t-1)=R$ is the fuzzy relation between $F(t-1)$ and $F(t)$, and "o" is the min-max operator. Since

R(t,t-1) is independent of time t, and only two consecutive times are involved in the model, it is called the first-order time invariant model. In forecasting applications, one might encounter two situations: One is when historical data are linguistic values and the other when data are numerical. In the second case, the data need to be fuzzified first in order to apply fuzzy time series models. In the following section, a step-by-step forecasting procedure will be introduced to indicate how to set up and utilize a first-order time-invariant model to forecast enrollments. Specific example data from The University of Alabama will be given to make the procedure more understandable.

2.2 Procedures of using fuzzy time series to forecast

Step 1. Define a range U on which the historical data are and upon which the fuzzy sets will be defined. Usually, when defining U, you must find first the minimum enrollment D_{min} and the maximum enrollment D_{max} of known historical data. Based on D_{min} and D_{max}, define U as $[D_{min}-D_1, D_{max}+D_2]$ where D_1 and D_2 are two proper positive numbers. In our example, we have on hand the enrollment data of the university from 1971 to 1990 with $D_{min}=13055$ and $D_{max}=19328$. For simplicity, we choose $D_1=55$ and $D_2=672$. Thus, the range of the interval U=[13000,20000].

Step 2. Partition U into several equally lengthy intervals. In our example, we divide U into 7 intervals of equal length. We use u_1, u_2, u_3, u_4, u_5, u_6 and u_7 for each interval, i.e., $u_1=[13000, 14000]$, $u_2=[14000, 15000]$, $u_3=[15000, 16000]$, $u_4=[16000, 17000]$, $u_5=[17000, 18000]$, $u_6=[18000, 19000]$ and $u_7=[19000, 20000]$.

Step 3. Define fuzzy sets on U. First, determine some linguistic values. For the linguistic variable "enrollment", let A_1=(not many), A_2=(not too many), A_3=(many), A_4=(many many), A_5=(very many), A_6=(too many), and A_7=(too many many) be the possible values. There is no restriction on the number of fuzzy sets defined. Second, define fuzzy sets on U. All the fuzzy sets will be labeled by the possible linguistic values. In our example, u_1, u_2, ..., and u_7 are chosen as the elements of each fuzzy set. To determine the memberships of u_1, u_2,......., and u_7 to each A_i (i=1 to 7), make a judgment of how well each u_k (k=1 to 7) belongs to A_i. If a u_k belongs to A_i completely, the membership will be 1; if u_k does not belong to A_i at all, the membership will be 0; otherwise, choose a number from (0,1) as the degree to which u_k belongs to A_i. From our experience, we have determined the membership for each element in the respective fuzzy sets. Thus, all the fuzzy sets A_i (i=1 to 7) are expressed as follows:

$$
\begin{aligned}
A_1 &= \{u_1/1,\ u_2/.5,\ u_3/0,\ u_4/0,\ u_5/0,\ u_6/0,\ u_7/0\};\\
A_2 &= \{u_1/.5,\ u_2/1,\ u_3/.5,\ u_4/0,\ u_5/0,\ u_6/0,\ u_7/0\};\\
A_3 &= \{u_1/0,\ u_2/.5,\ u_3/1,\ u_4/.5,\ u_5/0,\ u_6/0,\ u_7/0\};\\
A_4 &= \{u_1/0,\ u_2/0,\ u_3/.5,\ u_4/1,\ u_5/.5,\ u_6/0,\ u_7/0\}; \qquad (2)\\
A_5 &= \{u_1/0,\ u_2/0,\ u_3/0,\ u_4/.5,\ u_5/1,\ u_6/.5,\ u_7/0\};\\
A_6 &= \{u_1/0,\ u_2/0,\ u_3/0,\ u_4/0,\ u_5/.5,\ u_6/1,\ u_7/.5\};\\
A_7 &= \{u_1/0,\ u_2/0,\ u_3/0,\ u_4/0,\ u_5/0,\ u_6/.5,\ u_7/1\}.
\end{aligned}
$$

where u_i (i=1 to 7) is the element and the number below "/" is the membership of u_i to A_j (j=1 to 7).

For simplicity, we will also use A_1, A_2, ... , and A_7 as row vectors whose elements are the corresponding memberships in (2).

Step 4. Fuzzify the historical data, i.e., finding an equivalent fuzzy set for each year's enrollment. To do this, the degree of each year's enrollment belonging to each A_i (i=1 to 7) is determined. The process is the same as that of determining the memberships of u_i to A_j in Step 3. The equivalent fuzzy sets for each year's enrollment are shown in Table 1 with each fuzzy set containing seven elements.

Step 5. Obtain the fuzzy logic relations from Table 1 about the evolution of the enrollment of this university to set up the forecasting model. To do so, assume that if the maximum membership of one year's enrollment is under A_k, then we treat this year's enrollment as A_k. For example, for 1982, the maximum membership is under A_3, then we say that the enrollment of 1982 is A_3, or many. We can do the same for the rest. Thus, we can transform historical data into linguistic values. Since we are looking for the laws governing any two successive years' enrollments, in terms of fuzzy sets and fuzzy conditional statements, we will develop such logical relationships as "If the enrollment of year i is A_k then that of year i+1 is A_j", and so on. Thus, we can obtain all the fuzzy logic relations as follows (Note: the repeated relationships are counted only once):

$A_1 \text{-->} A_1$, $A_1 \text{-->} A_2$, $A_2 \text{-->} A_3$, $A_3 \text{-->} A_3$, $A_3 \text{-->} A_4$, $A_4 \text{-->} A_4$, $A_4 \text{-->} A_3$, $A_4 \text{-->} A_6$, $A_6 \text{-->} A_6$ and $A_6 \text{-->} A_7$.

Next, let us define an operator "x" of two vectors. Suppose C and B are row vectors of dimension m and $D=(d_{ij})=C^T x B$. Then the element d_{ij} of matrix D at row i and column j is defined as $d_{ij}=\min(C_i,B_j)$ (i,j=1 to m) where C_i and B_j are the i^{th} and the j^{th} elements of C and B respectively.

Let $R_1=A^T{}_1 x A_1$, $R_2=A^T{}_1 x A_2$, $R_3=A^T{}_2 x A_3$, $R_4=A^T{}_3 x A_3$, $R_5=A^T{}_3 x A_4$, $R_6=A^T{}_4 x A_4$, $R_7=A^T{}_4 x A_3$, $R_8=A^T{}_4 x A_6$, $R_9=A^T{}_6 x A_6$ and $R_{10}=A^T{}_6 x A_7$. Then, we get

$$R(t,t\text{-}1)= R = \bigcup_{i=1}^{10} R_i \qquad (3)$$

where R is a 7x7 matrix and "U" is the union operator. Using formula (3), some calculation yields:

$$R=\begin{bmatrix} 1 & 1 & .5 & .5 & 0 & 0 & 0 \\ .5 & .5 & 1 & .5 & .5 & 0 & 0 \\ 0 & .5 & 1 & 1 & .5 & .5 & .5 \\ 0 & .5 & 1 & 1 & .5 & 1 & .5 \\ 0 & .5 & .5 & .5 & .5 & .5 & .5 \\ 0 & 0 & 0 & 0 & .5 & 1 & 1 \\ 0 & 0 & 0 & 0 & .5 & .5 & .5 \end{bmatrix}$$

Thus, the forecasting model is:

$$A_i=A_{i-1} o R \qquad (4)$$

where A_{i-1} is the enrollment of year i-1 and A_i the forecasted enrollment of year i in terms of fuzzy sets.

Step 6. Calculate the forecasted outputs. Suppose the enrollment of year t is known and can be found from Table 1, to forecast the enrollment of year t+1, let A_{i-1} in (4) be the enrollment at year t and apply formula (4). Then A_i will be the forecasted enrollment of year t+1. For 1972 to 1991, the forecasted outputs are shown in Table 2.

Step 7. Interpret the forecasted outputs. The calculation results of formula (4) are all fuzzy sets. If the results in the form of fuzzy sets can satisfy the requirement for forecasting, just stop here. But in many cases, an equivalent scalar is desired. Therefore, translating the fuzzy output into a regular number is indeed a necessary step. Sometimes, this step is called defuzzification (Kosko,1992). For defuzzification in this procedure, we suggest to use the following steps:

1). If the membership of an output has only one maximum, then select the midpoint of the interval corresponding to the maximum as the forecasted value;

2). If the membership of an output has two or more consecutive maximums, then select the midpoint of the corresponding conjunct intervals as the forecasted value;

3). Otherwise, standardize the fuzzy output and use the midpoint of each interval to calculate the centroid of the fuzzy set as the forecasted value.

Following the above steps, we have obtained predicted values for enrollments from 1972 to 1991. The results are listed in Table 2 and shown in Fig.1 where the solid line is the actual enrollment and the dashed line is the forecasted enrollment. Note that we did not use enrollment data of 1991 to develop the forecasting model. The forecasting errors range from 0.1% to 8.7% with the average error being 3.18%. For 1991, the forecasting error is 1.7%. For a mid-term forecasting model, an average error of 3.18% is quite satisfactory.

3. The Artificial Neural Network Model

3.1 A Brief Introduction to Artificial Neural Networks

Technically, an artificial neural network is composed of a set of artificial neurons. Each neuron is a computational unit and has some inputs and output(s). Generally, the output of the neuron is a

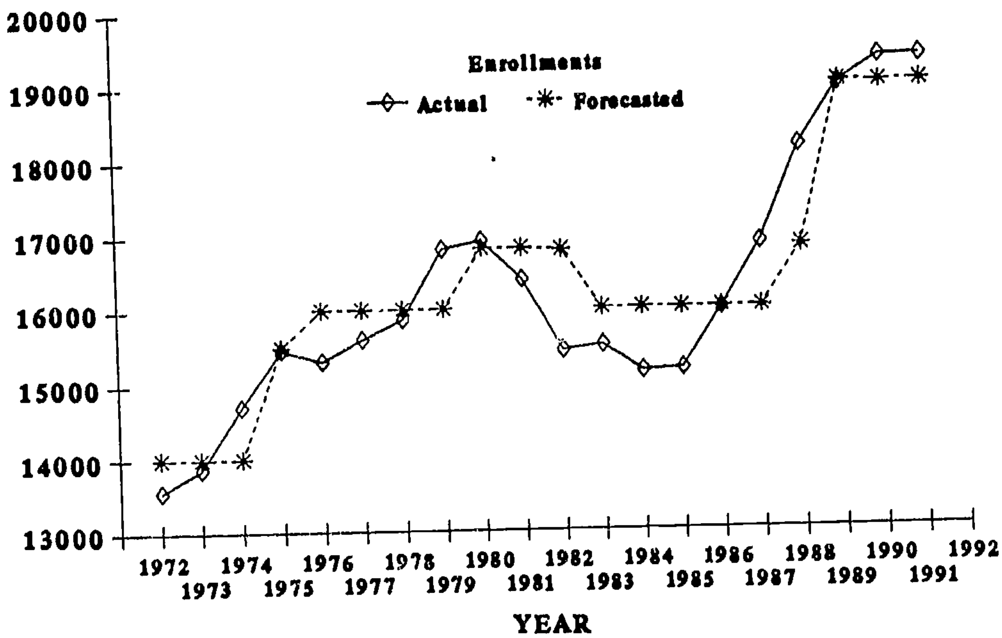

Fig.1 Forecasting results with fuzzy time series

nonlinear function of the weighted sum of all the inputs. There are connections among some of the neurons, and because of this, neural networks can process informati‸‸ in parallel. By connectionism and parallelism, neural networks have gained remarkable computational powers in som‸ areas that digital computers cannot match. In engineering, neural networks have been successf‸lly applied in signal processing, pattern recognition, phonem‸ recognition and many other areas (Widrow et al, 1988; Freeman et al, 1991; Kong et al, 1991; Kosko, 1992). Of all the various networks, the simplest and most fundamental might be the 3-layer backpropagation network. A typical 3-layer backpropagation network has the following structure:

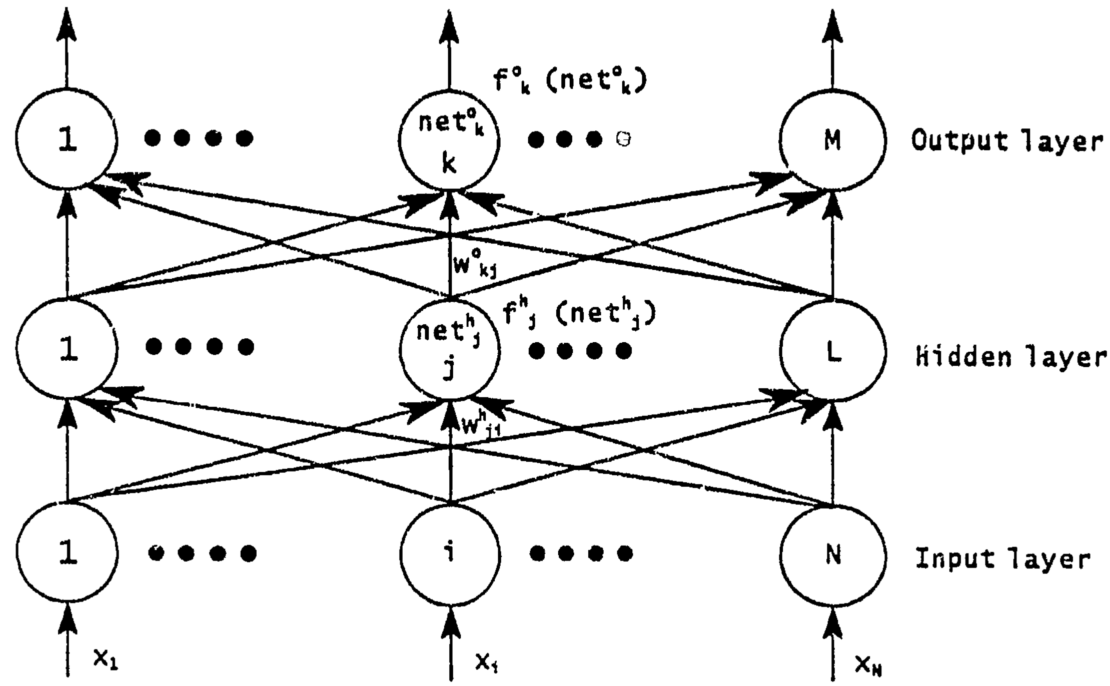

Fig.2 A typical structure of 3-layer neural networks

In the structure shown in Fig.2, the first layer is called the input layer, the second the hidden layer and the last the output layer. The input of the network is fed into each neuron of the input layer. The output of the neuron is the same as the input. The output of the input neuron, multiplied by the weight of the connection between the input neuron and the hidden neuron, is fed into the neurons of the hidden layer. Mostly, the output of hidden neuron j is a function of the weighted sum of all the inputs to the neuron and the function has the following form:

$$f^h_j(net^h_j) = (1 + \exp(-net^h_j))^{-1} \quad j = 1 \text{ to } L \qquad (5)$$

where $net^h_j = \sum_{i=1}^{N} w^h_{ji} * x_i$, w^h_{ji} is the weight of the connection from input neuron i to hidden neuron j and x_i is the ith element of input X. Similarly, the output of hidden neurons, multiplied by the weight of the connection between hidden neurons and output neurons, is fed into the output neurons. The output of output neuron k is a function of the weighted sum of all the inputs to the neuron. Normally, the function has the form of (6):

$$f^o_k(net^o_k) = (1+\exp(-net^o_k))^{-1} \quad k=1 \text{ to } M \qquad (6)$$

where $net^o_k = \sum_{j=1}^{L} w^o_{kj} * f^h_j(net^h_j)$ and w^o_{kj} is the weight of the connection from hidden neuron j to output neuron k.

The most important characteristic of neural networks is that they are trained instead of programmed. They have the ability to generalize once they have been trained with some exemplars. When presented with the input and the desired output called an exemplar at the input layer and the output layer respectively, the network will produce an output. If the output is compared with the desired output, an error can be determined. To improve the performance of the neural network, the weights w^h_{ji} and w^o_{kj} can be adjusted according to some rules (Freeman et al, 1991) such that the mean square of the error will be minimal. In neural networks literature, this process is called learning or training. Once a network has been trained, it has the ability to recall the desired output if the corresponding input is presented. In practice, a network can learn more than one exemplar simultaneously, and because of this ability, neural networks have found wide application.

3.2 Forecasting Scheme I: Using Original Enrollment Data

When applying neural networks for forecasting enrollments, there are three practical problems that should be considered. The first is the selection of the input, the second is how many input neurons we should choose for the network, and last is the problem of scaling the data.

As discussed above, many factors have influences on enrollments and all these factors can be taken as the input of the neural network. But, if we take too many influencing factors into account, the access to their sources may be a problem. The simplest case might be such that only current and past enrollment data are chosen as the influencing factor. The advantage of this is the ease with which to obtain these data. Thus, the input to the neural network will be present and past enrollment data. Obviously, the output from the network will be the forecasted enrollment.

Since only one factor, current and past enrollments, is considered, to make the forecasting errors as small as possible we shall use several years' enrollments as the input. In this study, we will use the current and the past 9 years' enrollments as input to forecast the next year's enrollment.

Since the output of a 3-layer back-propagation network is somewhere between 0 and 1 and the enrollments are from 0 to a large number, we first transform the enrollments into a number between 0 and 1 by dividing the enrollment by a proper positive number (100,000 in this study). Hereafter, the enrollments will be the scaled values unless otherwise noted.

When enrollment forecasting is compared with digital signal prediction in engineering, we find that these two processes share a common characteristic, both of them extrapolate the future by using present and historical information. Thus, some techniques applied in digital signal prediction can be borrowed for enrollment forecasting.

In signal prediction, the adaptive filter might be the simplest and the most widely used technique. A typical scheme of adaptive filters can be found in Widrow and Winter (1988). Nevertheless, it should be pointed out that adaptive filters are different from neural networks conceptually. To apply neural networks for enrollment forecasting, we shall adopt the scheme of adaptive filters with some modification. Instead of the adaptive filter, a 3-layer back-propagation neural network will be used. The scheme is shown in Fig.3.

In Fig.3, the neural network has 10 input neurons, 5 hidden neurons and only 1 output neuron.

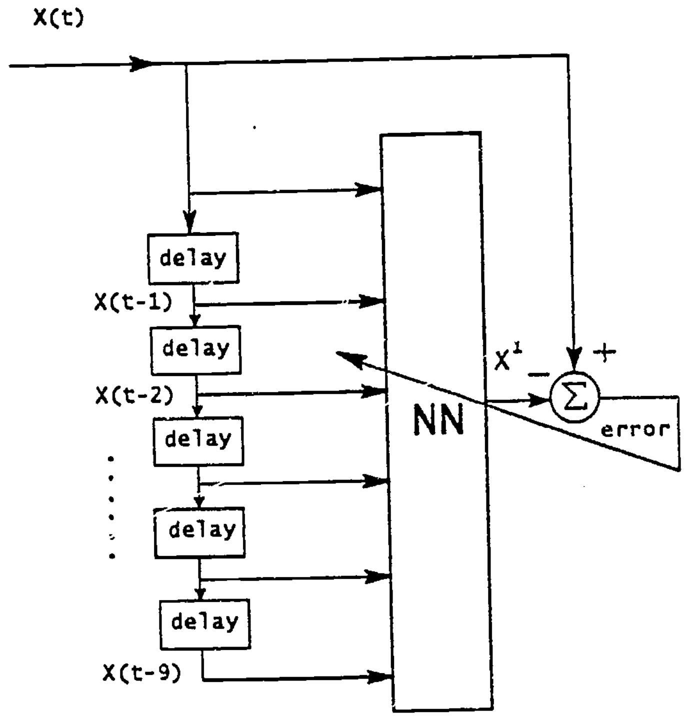

Fig.3 Forecasting enrollments with neural networks: Scheme 1

The procedure of using neural networks for forecasting enrollments is as follows:

Suppose we have the enrollment data of year t and previous years. To forecast the enrollment of year t+1, X(t+1), the neural network is trained using the following two exemplars:

$(Input_1; Output_1)=(X(t\text{-}10),X(t\text{-}9),X(t\text{-}8),X(t\text{-}7),X(t\text{-}6),X(t\text{-}5),X(t\text{-}4),X(t\text{-}3),X(t\text{-}2),X(t\text{-}1);X(t))$

$(Input_2; Output_2)=(X(t\text{-}11),X(t\text{-}10),X(t\text{-}9),X(t\text{-}8),X(t\text{-}7),X(t\text{-}6),X(t\text{-}5),X(t\text{-}4),X(t\text{-}3),X(t\text{-}2);X(t\text{-}1))$

where X(t) is the enrollment at year t.

After training the network, the following input data are fed into the network:

(X(t-9),X(t-8),X(t-7),X(t-6),X(t-5),X(t-4), X(t-3), X(t-2),X(t-1),X(t)).

The output of the network will be taken as the forecasted enrollment of year t+1, X(t+1). Multiply X(t+1) by 100,000 and we will get the actual forecasted enrollment of year t+1.

Following the above procedure, we have forecasted enrollments from 1984 to 1992. The neural network was simulated on the PlaNet neural network software package (Miyata, 1991). For 1984 to 1991, the forecasting errors range from 1.6% to 9.6% and the average forecasting error is 5.2%. The results are shown in Fig.4 where the solid line is actual enrollment and the dashed line forecasted enrollments. For mid-term enrollment forecasting, an average forecasting error of 5.2% is acceptable, because in this model we have taken into account only one influencing factor—the present and the past enrollments. But, there is still the potential for improvement of the accuracy. The cause for this error might be that when using the historical data, the neural network cannot recognize the trend of the evolution of the enrollment to forecast the future very well. To improve this, we shall derive a simple equation for forecasting enrollments in the next section.

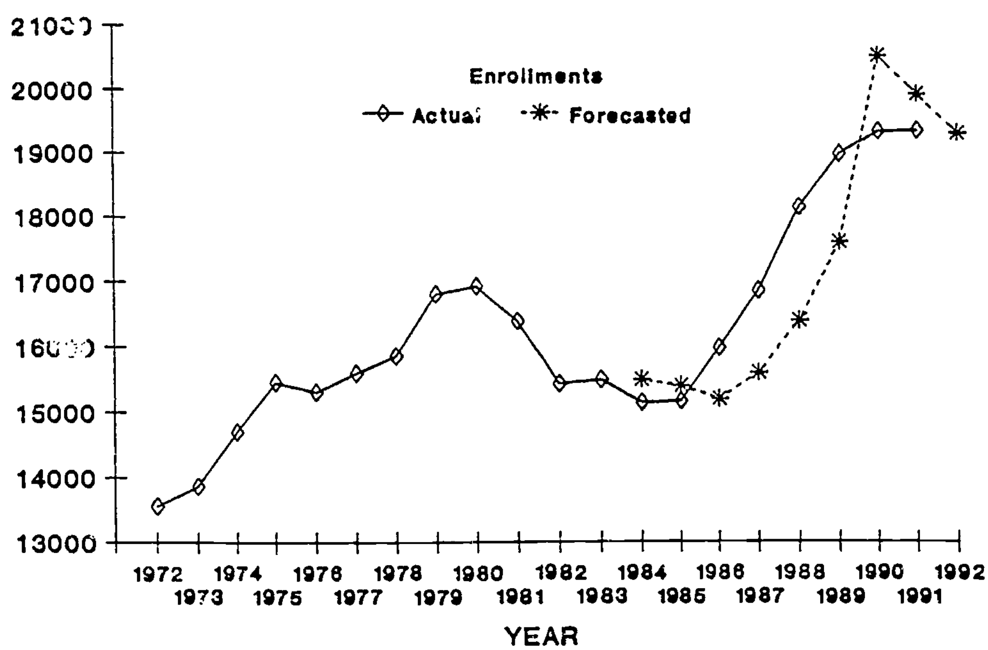

Fig.4 Forecasting results with Scheme 1

3.3 Forecasting Scheme II: Using Derivatives of Enrollments

Suppose the enrollment $X(t)$ is a differentiable function of time t. Given t_o and t where $t>t_o$, according to Taylor's Theorem (Lightstone, 1965), the enrollment $X(t)$ can be approximately expressed as follows:

$$X(t) \approx X(t_o)+X'(t_o)*(t-t_o) \quad (7)$$

where $X'(t_o)$ is the derivative of $X(t)$ with respect to time t at t_o. By definition,

$$X'(t_o)= \lim_{t->t_o} (X(t)-X(t_o))/(t-t_o)$$

In our case, since we are forecasting next year's enrollment, let $t=t_o+1$. Thus,

$$X'(t_o) \approx (X(t)-X(t-1))/(t-t_o) \quad (8)$$

$$\approx X(t)-X(t-1).$$

Therefore,

$$X(t) \approx X(t-1)+(X(t)-X(t-1))$$

$$\approx X(t-1)+\underline{DX}(t) \quad (9)$$

where $\underline{DX}(t)=X(t)-X(t-1)$.

Since $X(t)$s are historical data, Equation (9) implies that the enrollment at time t is approximately the sum of the enrollment at t-1 and the difference between the enrollments at t and t-1. This difference stands for the changing trend of the enrollments. From historical data, we can calculate $\underline{DX}(t)$ for each year. This tells us that if we forecast $\underline{DX}(t)$ instead of $X(t)$ using neural networks, the forecasting results might be improved. To test this hypothesis, we must first use historical enrollment $X(t)$ to calculate $\underline{DX}(t)$, and to obtain training exemplars for the network. Then, transform $\underline{DX}$'s into a number within [0,1] so that they can be applied to the network. The transformation formula is as follows:

$$DX(t)=(\underline{DX}(t)+D)/N \quad (10)$$

where $\underline{DX}(t)$ is the actual value, $D=|\min_t(\underline{DX}(t)|$ and N is a proper positive number (N is 1,000 in this study) such that $DX(t)$ will be within the interval [0,1]. In the following, we shall use DX for the transformed values. Finally, we will train the same network as used in the former section. The process is still the same.

Suppose we have transformed the enrollment data of year t and the previous years into DX's, to forecast DX(t+1), we should first use the following exemplars to train the network:

$(Input_1; Output_1)$=(DX(t-10),DX(t-9),DX(t-8),DX(t-7),DX(t-6),DX(t-5),DX(t-4),DX(t-3), DX(t-2),DX(t-1); DX(t))

$(Input_2; Output_2)$=(DX(t-11),DX(t-10),DX(t-9),DX(t-8),DX(t-7),DX(t-6),DX(t-5),DX(t-4), DX(t-3),DX(t-2); DX(t-1)).

And then input the following data into the network:

(DX(t-9),DX(t-8),DX(t-7),DX(t-6),DX(t-5),DX(t-4), DX(t-3),DX(t-2),DX(t-1),DX(t)).

The output of the network will be DX(t+1). To get the actual forecasted $\underline{DX}$(t+1), use the following formula:

$$\underline{DX}(t+1)=N*DX(t+1)-D \qquad (11)$$

where N is 1,000 in this study. Using equation (9), the forecasted enrollment X(t+1) will be:

$$X(t+1)=X(t)+\underline{DX}(t+1). \qquad (12)$$

Fig.5 is the scheme of using DX's to forecast enrollments.

Following the above steps, we have forecasted the DX(t)'s for each year from 1985 to 1992 and applied Equations (11) and (12) to calculate the forecasted enrollment for each year. The forecasted results are shown in Fig.6 where the solid line is actual enrollments and the dashed line forecasted enrollments. For 1985 to 1991, we have achieved forecasting errors ranging from 0.6% to 4.7% with the average error being 2.6%. This result is considered to be very satisfactory.

Compared with Scheme I, i.e., the method using only the original enrollment data, the accuracy of using $\underline{DX}$(t)'s has been doubled (the average error has been reduced from 5.2% to 2.6%). Very interestingly, the forecasted enrollment of 1992 will be less than the actual enrollment of 1991.

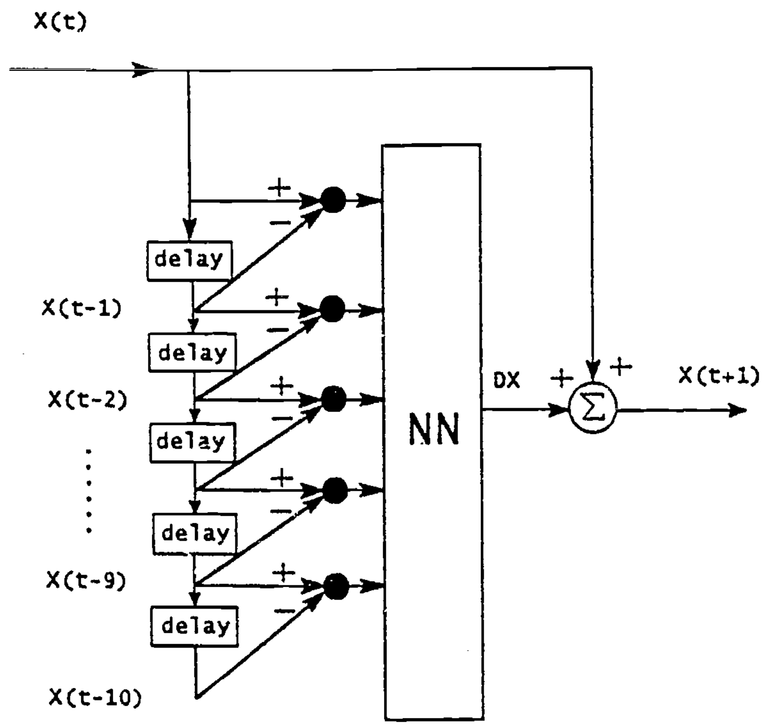

Fig.5 Forecasting enrollments with neual networks: Scheme 2

4. Discussions and Concluding Remarks

4.1. About the fuzzy time series model

Compared with any other forecasting methods mentioned in Section I, the method applied here has at least the following 5 advantages:

1. The average forecasting error of this method is smaller and acceptable.

2. Human experience knowledge can be utilized from the very beginning until the end of the forecasting process. If historical data cannot be obtained but the knowledge of the evolution of the

university's enrollments of the past can, i.e., the evolution laws in the form of "IF....THEN...." can be obtained, none of the other methods can be applicable. Nevertheless, in this case with fuzzy time series, we are still able to establish a forecasting model and make good forecasts. Our method is better than the others in this aspect;

3. Historical data have less important roles in our method. Note that in Section III, the data were only used in Step 4 of the procedure where the data were fuzzified. Once fuzzified, they were

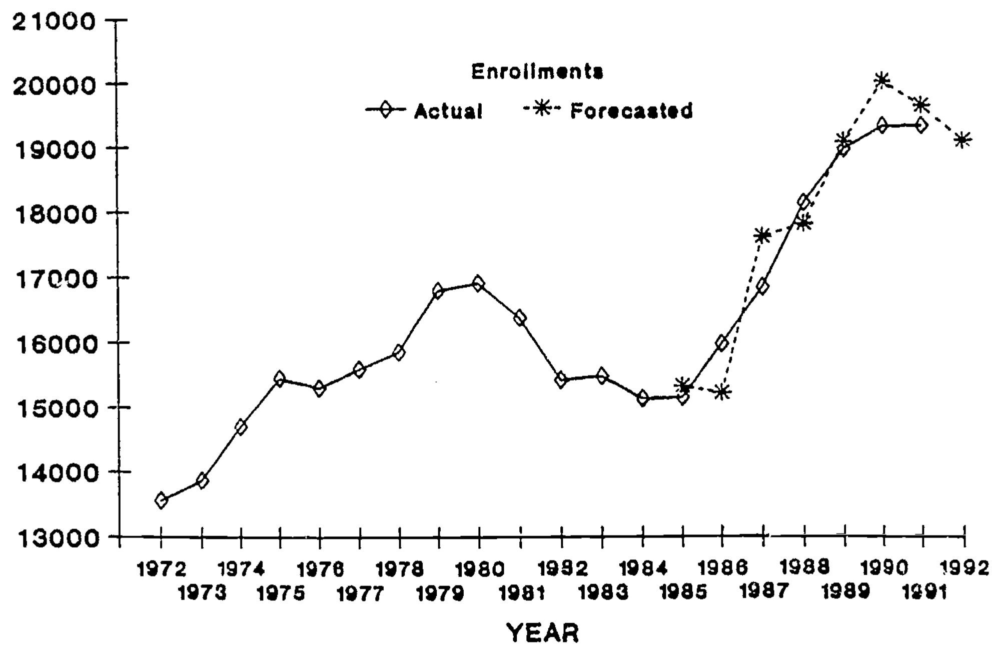

Fig.6 Forecasting results with Scheme 2

never used again. Although in Step 1, the minimum and the maximum of the historical data were used to define the universe, it is not necessary to do so. In effect, we could define a universe with only our experience. The principle is that the universe should cover the possible range of the enrollments;

4. Even if historical data are real numbers, since the model has robustness, the requirement for historical data is not very strict;

5. When historical data are linguistic values, the procedure is almost the same for forecasting

enrollments except that Step 4 is omitted and Table 1 contains the corresponding fuzzy sets of (2) as row vectors.

The major problem with this method is that the forecasted values depend to some degree on our interpretations of the outputs of the forecasting model in Step 7. Different interpretations may lead to different forecasted results. This makes the process quite subjective. To overcome this shortcoming, an objective method should be applied.

4.2. About the neural network models

We have used artificial neural networks as a mid-term model to forecast enrollments of the University of Alabama. Two schemes were applied in which the improved scheme was better than the original. This indicates that when applying neural networks for forecasting enrollments, we need to modify the scheme of adaptive filters. The final forecasted results of Scheme II are very satisfactory. Since only one factor, the present and the past enrollment information, has been considered in the model and very accurate forecasting results have been achieved, neural networks Scheme 2 is a competitive model for forecasting.

Also, the results show that using the original data in neural networks for forecasting enrollments might not be suitable since the neural network cannot recognize the trend of the enrollments very well. To make up this deficiency, DX's were included in the model and the neural network model actually forecasts DX's. We prefer to use Scheme II. Obviously, the neural network scheme in Fig.5 can be applied to many other areas.

Compared to other models discussed previously, using neural networks to forecast enrollments has at least the following two major advantages:

(1). The model can be easily programmed in any programming language and run on a PC. When using this model to forecast enrollments, the administrators in the Admissions Office only need to input the current and past nine years' enrollment data, and the model can produce next year's forecasted enrollment. The administrators need not be involved with the computation process as is necessary in many other models.

(2). There are no factors that need to be determined by human beings in advance and this makes the forecasting process objective.

Acknowledgement. The material on neural network forecast model was primarily based on a class project by one of the authors in EE587 at the University of Alabama. We acknowledge our sincere thanks to Dr. R. E. Smith, the instructor of that course.

References

Chatman, S. P. (1986), Short-term forecasts of the number and scholastic ability of enrolling freshman by academic divisions. *Res. Higher Educ.* 25(1): 68-81.

Freeman, J. A. & Skapura, D. M. (1991). *Neural networks: algorithms, applications, and programming techniques*. Addison-Wesley Publishing Company, Inc.

Gardner, D. E. (1981). Weight factor selection in double exponential smoothing enrollment forecasts. *Res. Higher Education.* 14(1): 49-56.

Hoenack, S. A. & Weiler, W. C. (1979). The demand for higher education and institutional enrollment forecasting. *Economic Inquiry*, 17: 89-113.

Kong, S. & Kosko, B. (1991). Differential competitive learning for centroid estimation and phoneme recognition. *IEEE Trans. Neural Networks*, 2(1): 118-124.

Kosko, B. (1992). *Neural networks and fuzzy systems*, Prentice Hall.

Lightstone, A. H. (1965). *Concepts of calculus*. Harper & Row Publishers, New York, Evanston, and London.

Miyata, Y. (1991). *A user's guide to PlaNet Version 5.6.*

Paulsen, M. B. (1989). A practical model for forecasting new freshmen enrollment during the application period. *Coll. Univ.* 64(4): 379-91.

Pope, J. A., & Evans, J. P. (1985). A forecasting system for college admissions. *College and University* 60(1): 113-131.

Shaw, R. C. (1984). Enrollment forecasting: what methods work best? *NASSP Bull* 68: 52-8.

Song, Q. and Chissom, B. S. (1993a). Fuzzy time series and its models, *Fuzzy Sets and Systems* 54(3): 269-278.

Song, Q., and Chissom, B. S. (1993b). Forecasting enrollments with fuzzy time series: Part I. *Fuzzy*

Sets and Systems 54(1): 1-10.

Warrack, B. J. & Russell, C. N. (1983). Forecasting demand for postsecondary education in Manitoba: The motivational index and the demand index as an enrollment forecasting tool. *Res. Higher Educ.* 19(3): 335-49.

Weiler, W. C. (1980). A model for short-term institutional enrollment forecasting. *J. of Higher Education*, 51(3): 314-327.

Widrow, B. & Winter R. (1988). Neural nets for adaptive filtering and adaptive pattern recognition. *Computer*, 21(3): 25-39.

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

L.A. Zadeh (1975a), Fuzzy logic and approximate reasoning, *Synthese*, 30: 407-428.

L.A. Zadeh (1975b), The concept of a linguistic variable and its application to approximate reasoning: Part 1, *Inform. Sci.*, 8: 199-249;

L.A. Zadeh (1975c), The concept of a linguistic variable and its application to approximate reasoning: Part 2, *Inform. Sci.*, 8: 301-357;

L.A. Zadeh (1975d), The concept of a linguistic variable and its application to approximate reasoning: Part 3, *Inform. Sci.*, 9: 43-80.

Table 1

	A1	A2	A3	A4	A5	A6	A7
1990	0	0	0	.3	.5	.8	1
1989	0	0	0	.25	.55	1	.8
1988	0	0	.1	.5	.8	1	.7
1987	0	.1	.5	1	.8	.1	0
1986	0	.2	1	.7	.2	0	0
1985	.2	.8	1	.2	0	0	0
1984	.2	.8	1	.2	0	0	0
1983	.2	.8	1	.2	0	0	0
1982	.2	.8	1	.2	0	0	0
1981	0	.2	.8	1	.5	0	0
1980	0	.1	.5	1	.9	.2	0
1979	0	.1	.5	1	.9	.2	0
1978	0	.5	1	.7	.2	0	0
1977	0	.6	1	.6	.1	0	0
1976	.2	.8	1	.2	0	0	0
1975	.2	.8	1	.2	0	0	0
1974	.8	1	.8	.1	0	0	0
1973	1	.9	.2	0	0	0	0
1972	1	.8	.1	0	0	0	0
1971	1	.5	0	0	0	0	0

Table 2

year	output membership	standardized membership	predicted value
1972	1,1,.5,.5,.5,0,0	.286, .286, .143, .143, .143, 0, 0	14000
1973	1,1,.8,.5,.5,.1,.1	.25, .25, .2, .125, .125, .025, .025	14000
1974	1,1,.9,.5,.5,.2,.2	.2325, .2325, .209, .116, .116, .047, .047	14000
1975	.8,.8,1,.8,.5,.5,.5	.163,.163, .204, .163, .102, .102, .102	15500
1976	.5, .5, 1, 1, .5, .5, .5	.111, .111, .222, .222, .111, .111. .111	16000
1977	.5, .5, 1, 1, .5,.5,.5	.111, .111, .222, .222, .111, .111, .111	16000
1978	.5, .5, 1, 1, .5, .6, .5	.109, .109, .217, .217, .109, .13, .109	16000
1979	.5, .5, 1, 1, .5, .7, .5	.106, .106, .213, .213, .106, .149, .106	16000
1980	.1, .5, 1, 1, .5, 1, .5	.0217, .108, .217, .217, .108, .217, .108	16813
1981	.1, .5, 1, 1, .5, 1, .5	.0217, .108, .217, .217, .108, .217, .108	16813
1982	.2, .5, 1, 1, .5, 1, .5	.0425, .106, .213, .213, .106, .213, .106	16789
1983	.5, .5, 1, 1, .5, .5, .5	.111, .111, .222, .222, .111, .111, .111	16000
1984	.5, .5, 1, 1, .5, .5, .5	.111, .111, .222, .222, .111, .111, .111	16000
1985	.5, .5, 1, 1, .5, .5, .5	.111, .111, .222, .222, .111, .111, .111	16000
1986	.5, .5, 1, 1, .5, .5, .5	.111, .111, .222, .222, .111, .111, .111	16000
1987	.2, .5, 1, 1, .5, .7, .5	.045, .114, .227, .227, .114, .159, .114	16000

1988	.1, .5, 1, 1, .5, 1, .5	.027, .108, .217, .217, .108, .217, .108	16813
1989	0, .5, .5, .5, .5, 1, 1	0, .125, .125, .125, .125, .25, .25	19000
1990	0, .5, .5, .5, .5, 1, 1	0, .125, .125, .125, .125, .25, .25	19000
1991	0, .5, .5, .5,. 5, .8, .8	0, .138, .138, .138, .138, .222, .222	19000