Process Identification Using Artificial Neural Network

by

Dr. Arshad Ahmad
Department of Chemical Engineering,
Universiti Teknologi Malaysia, Kuala Lumpur

ABSTRACT

Recent years has seen the emergence of a new paradigm in system's identification known as Artificial Neural Network (ANN). ANN is a methodology inspired from the structure and mechanism of human brain. Similar to the human brain (albeit in simplistic scale), ANN has the learning capability that enables it to remember past information. ANN is also known to be proficient in approximating nonlinear function to arbitrary accuracy in a black-box manner. As such, if adequate training over a sufficiently rich data is provided, the network will be able to capture the information contained within the data and store the in the form of model which can then be utilize to predict future characteristics.

This paper describes the basic mechanism of ANN and its application in the identification of a polymerization process. The results obtained highlights the proficiency of ANN models in predicting the reactor product concentration, thus recommending its application in other model-related process engineering tasks.

INTRODUCTION

A man accumulated knowledge through successive learning processes since he was born. All his judgment, decisions, views as well as actions are strongly influenced by what he has learned in the past plus current information which he manage to fit in one of the previous "models" within his reasoning system. The accuracy of his decision or judgment depends greatly on how well the current situation that he is dealing with, fit the model consisting of information that he has learned. If he is confronted with a totally new situation, his evaluation on this unfamiliar condition will be based on extrapolation of the existing model. The outcome of this new experience which could be positive (*e.g.*, correct decision) or negative is then added to his knowledge.

The learning process of a man being takes place in his brain, a "grey matter" made up of many millions of highly inter-connected individual nerve cells called neurons. These neurons are capable of receiving and processing incoming information as well as transmitting the processed signals to other neurons. A simplified macroscopic explanation of the neuron's mechanism is as follows. External information from other neurons are received by the cell at connections called synapses. These synapses are connected to the cell inputs, or dendrites. When the total excitation from the dendrites exceeds a certain threshold, an electrical pulse is sent down the neuron output linkage, known as the axon. These signals are then captured by other neurons at their synapses.

What really happens within the brain is much more complicated. Brain cells receive signals from all parts of the human body and provide the corresponding responses to facilitate human body functions. Within the cells and their connections, there are also facilities to store previous information and based on these, experiences are gathered. These experiences are utilize in future information processing so that the required response can be determined more accurately. It is this learning ability that inspired the development of Artificial Neural Network (ANN) and encourages its utilization in widespread applications.

NETWORK STRUCTURE

ANNs are made up of individual models of the biological neuron that are connected together to form a network. Unlike the human brain which consists of between 10^{11} to 10^{14} neurons, the number of artificial neurons in an ANN is relatively very small. The connections within the ANN is also much more simplified than that of the brain. However, despite the simplicity, an ANN designed using only a few neurons can be trained to learn information such as patterns and process dynamics.

An ANN is constructed of interconnected basic elements called neurons or nodes. A schematic diagram of this neuron is shown in Fig.1. Similar to their biological counterpart, these neurons are capable of processing incoming information and transferring them to other neurons. The input signals come from either the environment or outputs of other neurons through connections as specified by the network architecture.

Associated with each connection is an adjustable value called the network weight. During learning, it is these weights that are adjusted to provide some kind of memory for the ANN. From a modelling point of view, these weights are analogous to model parameters which regulate the influence of each variable on overall model performance. Similarly, network weights serve as a measure for connection strength that control the influence of each incoming signal on the recipient neuron.

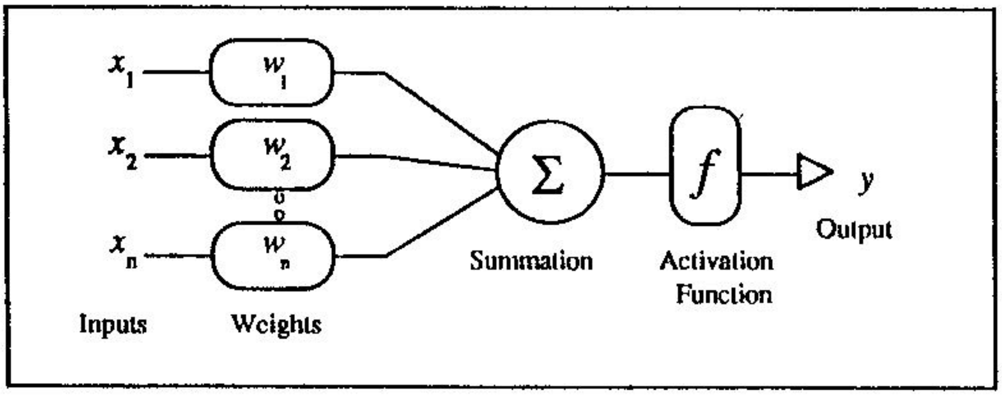

Fig.1 : Schematic Diagram of an Artificial Neuron

Within each neuron, input signals are summed and transformed using an activation function (also known as a threshold or squashing function) before being sent to other neurons. Transformation of data via activation functions is needed to impart pattern mapping ability to the networks. To enhance the identification ability, these neurons are arranged in a structure known as the network architecture. Currently, a wide selection of architecture have been developed but in the field of identification and control, the multilayer feedforward artificial neural network (FNN) has found prominent application. FNN is also referred to as multilayer perceptron (MLP) or back propagation networks in the literature. This network is constructed of neurons arranged in several layers. There is an input layer to receive the incoming data to the network, and an output layer to deliver the processed data from the network. In between these two layers, there could be several layers known as the *hidden layers*. A typical FNN architecture with two hidden layers is shown in Fig.2. The neurons in the input layer do not perform any processing and is used to distribute all inputs to the next layer. All other neurons carry out information processing as mentioned in the previous section. In addition to the regular neurons, bias neurons with fixed output of unity are often connected to all the neurons in the hidden and output layers.

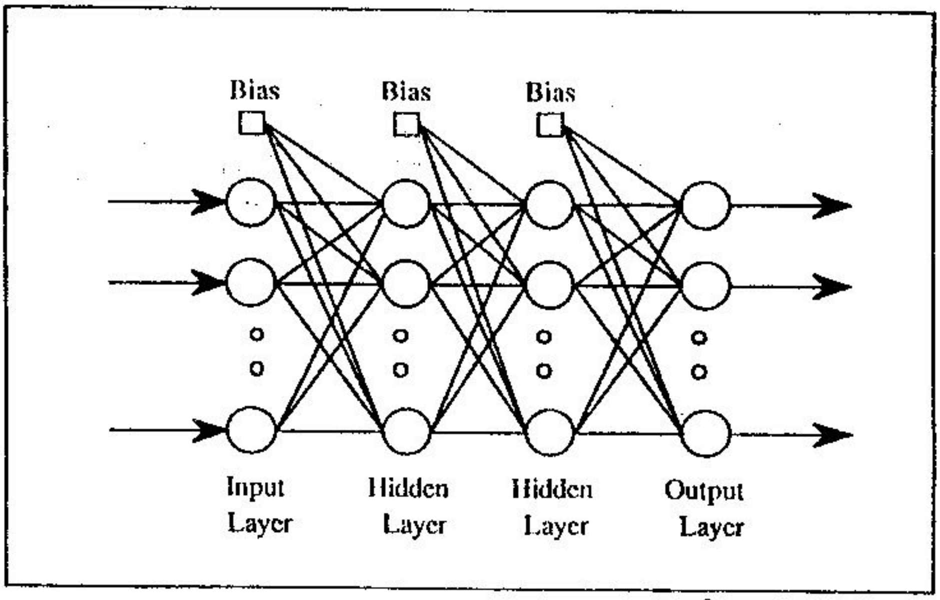

Fig.2: Topology of Multilayer Feedforward Network

It has been shown that a feedforward network with one hidden layer can serve as universal approximators (Cybenko, 1989). This implies that the network can learn the relationship between any set of inputs and outputs so that when given the inputs, the outputs can be reproduced. However, in order to obtain sufficiently good approximation qualities, a network with "sufficient" neurons must be trained with enough information so that the weights associated with all the connections within the network are optimize to achieve the desired input/output mapping. This can be achieved using a wide selection of learning algorithms. In process engineering applications, Back Propagation and its variants have been widely employed.

PREDICTION OF REACTOR DYNAMICS

To illustrate the modelling capabilities of ANN, a simulation studies involving a Polyethylene Terephthalate Transesterification reactor was carried out. The task was to predict the main product concentration namely, Bis-hydroxyethyl Terephthalate (BHET). In practice, this concentration can only be measured using off-line analysis in the laboratory, thus inflicting a large time delay. Since it is the primary aim the reactor operation to obtain BHET with high purity, frequent measurement of BHET is required for process control purposes.

A multiinput singleoutput model was constructed using feedforward network, trained using conjugate gradient algorithm. The inputs to the model were the current and previous values of the manipulated variable and the previous values of the output, while the output of the model is the future (predicted) value of the reactor concentration. Training data was generated from a deterministic reactor model published by Choi and Khan (1986) and validated using cross validation method. The prediction result is shown in Fig.3. The result confirms the proficiency of ANN in process identification and forecasting. Although this is based on a one-step ahead prediction, a multistep ahead prediction can also be devised using the same principles.

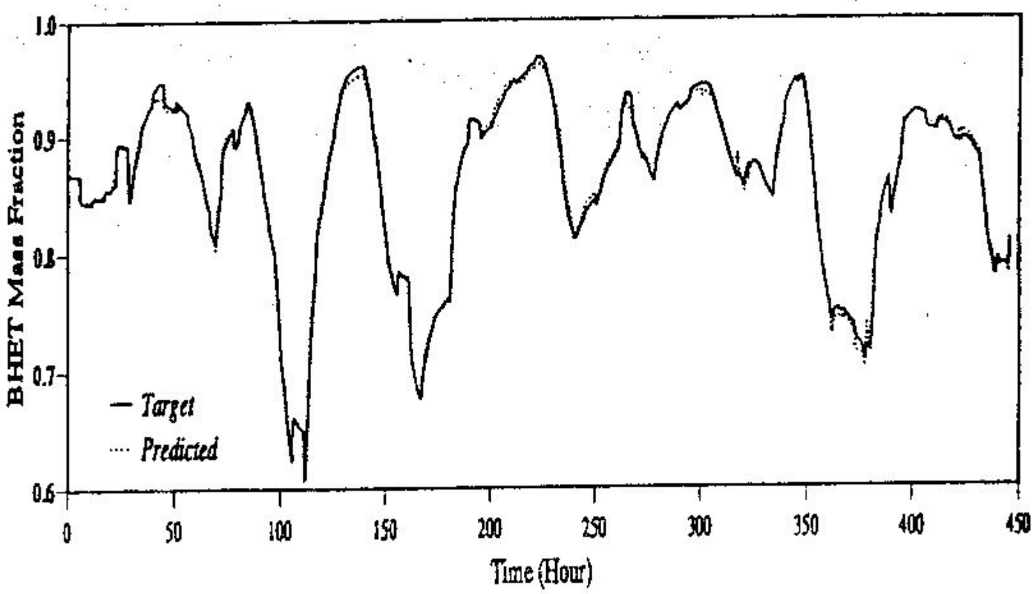

Fig.3 : Prediction for Reactor Concentration.

CONCLUSION

In this paper, the application of ANN models in process identification has been briefly described. The proficiency of ANN models in predicting a reactor concentration suggests that it can be used in process monitoring and control. However, ANN is not a panacea in nonlinear systems modelling. It is a powerful method and has indeed display promising identification capabilities, but it requires some time before it can be considered as a fully developed methodology.

REFERENCES

Choi K.Y. and A.Khan (1988). Optimal State Estimation in the Transesterification Stage of a Continuous Polyethylene Terephthalate Condensation Polymerization Reactors. *Chem Engng. Sci.*, **43**, No.4, 749-762

Cybenko, G. (1989). Approximation by Superpositions of Sigmoidal Function. *Math. Control Signals System*, **2**, 303-314.